时 间 序 列 分 析

谢衷洁 编著

北京大学出版社

内 容 提 要

本书的主要内容是时间序列分析的基础性知识，涉及到平稳随机过程、ARMA模型、马氏扩张、预测与滤波、模型拟合、谱估计及潜周期分析等，侧重于基本概念和方法的讲解，可作为进一步学习、研究时间序列分析的先导。

本书可供高等院校理工科高年级学生、研究生、大学教师及科技工作者学习参考。

时间序列分析

谢衷洁 编著

责任编辑：王明舟

北京大学出版社出版

（北京大学校内）

北京大学印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

850×1168毫米 32开本 11.5印张 298千字

1990年9月第一版 1990年9月第一次印刷

印数：00001—4,000册

ISBN 7-301-00849-X/O·154

定价：6.00元

前　言

本书是由1980年至1985年讲课的讲稿整理成的。讲课的对象是数学系、概率统计系高年级学生、研究生和进修教师。

本书的内容主要是时间序列分析的基础性知识，涉及到平稳随机过程，ARMA模型，马氏扩张，预测与滤波，模型拟合，谱估计及潜周期分析等。课程本身侧重于基本概念和方法的讲解而不着眼于某些过细分支内容的介绍，因此对学生讲它只起一个导引性的作用。

本书要求的先修课程是普通概率统计，线性代数，复变和实变函数论及泛函分析中有关Hilbert空间的一些知识。

本书有许多内容和思想是属于江泽培教授所领导的讨论班的，如极大一步预测误差准则下最优拟合的定理，弱P线性模型参数估计的强相容性及马氏扩张等等。

本书比较偏重于时间序列的概率分析而统计内容比较弱，应该说这是一个缺点。

感谢陈家鼎同志对本书出版的积极鼓励和支持，感谢李东风同志详细地阅读了全稿并提出了许多宝贵意见。

由于水平所限，错误一定不少，恳请读者批评指正。

谢衷洁

于北大燕东园

一九八六年十二月

目　录

第三章　时间序列的预报与滤波

第四章　谱估计的参数方法

第一章　平稳随机过程的基本知识

§1　随机过程的定义及例子

1°　随机过程的定义

在普通概率统计课程的学习中，我们已经学过一元、n元随机变量及其分布，并初步了解到运用这些理论和方法可以解决许多实际问题。但是有大量的课题涉及到的不是有限多个的随机变量，求解问题的出发点也不是n元随机变量的N个独立样本而是无穷多个随机变量的一次实现观测。这样，以往学过的知识就不足以解决这些问题，例如：

(1) 对某城市的气温进行n年的连续观测，得记录$\{x_t, a\leqslant t\leqslant b\}$，试问该地气温有没有以$T_0$年为周期的变化规律？

(2) 从杂乱电讯号的一段观测$\{y_t, 0<t<T\}$中，研究是否存在某种随机信号s_t？

(3) 监测器上收到某种电波记录$\{\xi_t, \alpha<t<\beta\}$，试问它是否确实是追踪的某个信号？

等等。

以上三个问题是随机过程论(特别是应用随机过程)中的重要问题，(1)属于谱分析，(2)属于过程检测，(3)属于过程识别。在这里我们首先遇到的是具有连续足标的无穷多个随机变量的一段观测样本，即随机过程的有限(或无限)观测样本。首先，我们需引入随机过程的定义：

定义1.1　设$(\Omega, \mathscr{F}, P)$为概率空间，T是某一足标集，若对任意的$t\in T$，$\xi_t(\omega)(\omega\in\Omega)$是该概率空间上的一个随机变量，则这一簇随机变量$\{\xi_t(\omega), t\in T\}$就称为是随机过程。

以后我们讲随机过程多指T是无穷多个元素，甚至主要是指

实数集 $\boldsymbol{R}_1$ 或直线上的一个区间 I；若 T 是整数集或非负整数集，就称$\{\xi_t(\omega)\}$为随机序列，特别，当 T 代表时间时就称为时间序列。如果不混淆就简记$\{\xi_t(\omega)\}$为 ξ_t 或 $\xi(t)$。ξ_t 大多指实值随机变量，但以后也要讨论复随机变量，即$\xi_t(\omega)=\eta_t(\omega)+\mathrm{i}\,\zeta_t(\omega)$，而 $\eta_t(\omega)$，$\zeta_t(\omega)$是实值随机变量。最简单的随机过程如

例1.1 设$(\Omega,\mathscr{F},P)$为对应于掷硬币的概率空间

$$\Omega=\{\text{正}=\omega_1,\ \text{反}=\omega_2\},\ \mathscr{F}=\{\omega_1,\omega_2,\varnothing,\Omega\},\quad P(\omega_1)=P(\omega_2)=\frac{1}{2},$$

作无穷多次独立实验可引入随机变量 $\xi_t(\omega)$：

$$\xi_t(\omega)=\begin{cases}0, & \omega=\omega_1,\\ 1, & \omega=\omega_2,\end{cases}\qquad t=0,1,2,\cdots. \tag{1.1}$$

则$\{\xi_t(\omega),\ t=0,1,2,\cdots\}$为随机序列。

设$\{\xi_t(\omega),t\in T\}$是随机过程，对 $t_1,t_2,\cdots,t_n\in T$，则$(\xi_{t_1}(\omega),\xi_{t_2}(\omega),\cdots,\xi_{t_n}(\omega))$是 n 元的随机变量。若 $T=\boldsymbol{R}_1$，或区间 I，则固定$\omega_0\in\Omega$，$\{\xi_t(\omega_0),t\in T\}$对应于一条曲线（见图1.1）。

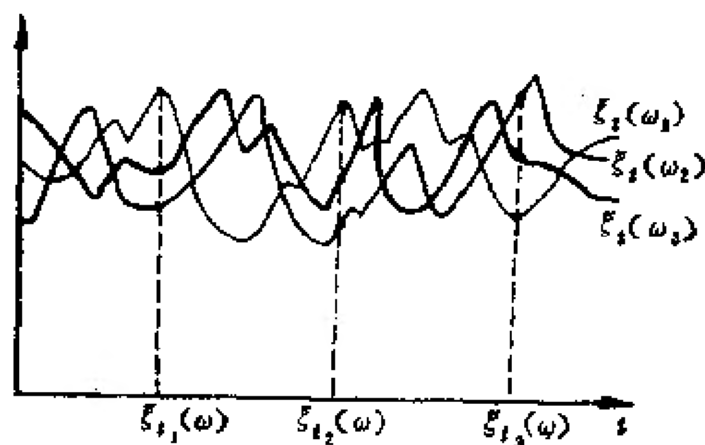


图 1.1

2° 随机过程的时律

设$\{\xi_t,t\in T\}$为某概率空间上的随机过程，对 $t_1,t_2,\cdots,t_n\in T$，$(\xi_{t_1},\cdots,\xi_{t_n})$为 n 元随机变量。由概率论知，此时必有一个 n 元分布函数

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n),\quad (x_1,x_2,\cdots,x_n)\in \boldsymbol{R}_n \qquad (1.2)$$

与之对应，考虑一切可能的 n 元分布函数的全体，记为 F_ξ，即

$$F_\xi=\{F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)\,|\,n>0,\ t_1,t_2,\cdots,t_n\in T\}. \qquad (1.3)$$

这一切有限维的分布函数簇 F_ξ 就称为是随机过程 ξ_t的时律。

由概率论知识不难看出，F_ξ 中的有限维分布函数 $F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)$具有以下两条重要性质：

(a) 对$(t_1,t_2,\cdots,t_n)$的任一排列 $(t_{\nu_1},t_{\nu_2},\cdots,t_{\nu_n})$，有

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=F_{t_{\nu_1},t_{\nu_2},\cdots,t_{\nu_n}}(x_{\nu_1},x_{\nu_2},\cdots,x_{\nu_n}). \qquad (1.4)$$

(b) $$F_{t_1,t_2,\cdots,t_m}(x_1,x_2,\cdots,x_m)=F_{t_1,\cdots,t_m,t_{m+1},\cdots,t_n}(x_1,\cdots,x_m,+\infty,\cdots,+\infty). \qquad (1.5)$$

以上两条件称为 Колмогоров 的相容性条件。

有穷维分布函数簇满足K氏相容性的条件，关系到以下的随机过程的存在性定理：

定理1.1 设给定一个足标集 T 以及一簇满足K氏相容性条件的有穷维分布函数簇

$$F=\{F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n),\quad n>0,\quad t_1,t_2,\cdots,t_n\in T\},$$

则必存在一个概率空间$(\Omega,\mathscr{F},P)$上的随机过程$\{\xi_t(\omega),t\in T\}$,使得其时律 $F_\xi=F$。

此定理在理论上具有重要意义。它肯定了对给定有穷维分布簇 F 必存在随机过程以 F 为时律，但需注意概率空间$(\Omega,\mathscr{F},P)$和$\{\xi_t(\omega)\}$的构造并不唯一。定理1.1的证明我们略去，有兴趣的读者可参看[1],[2]。

3° 正态过程及存在性定理

作为随机过程的重要例子，我们将介绍正态过程(或称Gauss

过程）及其存在性定理．首先介绍正态过程的定义：

定义1.2 设$\{\xi_t, t\in T\}$是实随机过程①，若对任意的n元向量$\boldsymbol{t}=(t_1,t_2,\cdots,t_n)$，$t_i\in T$，$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$的特征函数$f(\boldsymbol{u})$，$\boldsymbol{u}\in\boldsymbol{R}_n$，具有形式

$$f(\boldsymbol{u})=e^{i\boldsymbol{a}\boldsymbol{u}'-\frac{1}{2}\boldsymbol{u}\boldsymbol{\Sigma}\boldsymbol{u}'},\tag{1.6}$$

其中$\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$为实向量，$\boldsymbol{\Sigma}$为实对称非负定阵，则称$\{\xi_t\}$为实正态过程．

需说明的是：如果$\boldsymbol{\Sigma}>0$（正定阵），则(1.6)显然可以直接写出所对应的分布密度$N(\boldsymbol{a},\boldsymbol{\Sigma})$，其中$\boldsymbol{\Sigma}$的元素

$$\sigma_{ij}=E(\xi_{t_i}-a_i)(\xi_{t_j}-a_j),\qquad i,j=1,2,\cdots,n.$$

但如果$\boldsymbol{\Sigma}\geqslant 0$（非负定阵），即矩阵可能退化，这时虽然写不出$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$的概率密度，但(1.6)式的$f(\boldsymbol{u})$仍然是特征函数．事实上，可令

$$\boldsymbol{\Sigma}_N=\boldsymbol{\Sigma}+\frac{1}{N}\boldsymbol{I},\tag{1.7}$$

其中$\boldsymbol{I}$为单位阵，N为正整数，则对$\boldsymbol{\alpha}\in\boldsymbol{R}_n$，均有

$$\boldsymbol{\alpha}\boldsymbol{\Sigma}_N\boldsymbol{\alpha}'=\boldsymbol{\alpha}\boldsymbol{\Sigma}\boldsymbol{\alpha}'+\frac{1}{N}\boldsymbol{\alpha}\boldsymbol{\alpha}'\geqslant\frac{1}{N}\boldsymbol{\alpha}\boldsymbol{\alpha}'>0.$$

可见

$$f_N(\boldsymbol{u})=e^{i\boldsymbol{a}\boldsymbol{u}'-\frac{1}{2}\boldsymbol{u}\boldsymbol{\Sigma}_N\boldsymbol{u}'}=f(\boldsymbol{u})e^{-\frac{1}{2N}\boldsymbol{u}\boldsymbol{u}'}\tag{1.8}$$

为特征函数．

然而当$N\to\infty$时，在$\boldsymbol{u}$的任意有界集上，$f_N(\boldsymbol{u})$一致收敛到$f(\boldsymbol{u})$，又由于$f(\boldsymbol{u})$在0点是连续函数，由特征函数的极限定理知$f(\boldsymbol{u})$是特征函数．

① 全义讲应为“某概率空间上的取实值的随机过程”，以后多采用简化写法．

以下的 Gauss 过程存在定理对随机过程的理论和方法的研究具有重要意义。

定理1.2 设 T 为足标集，a_t 为任意实值函数，$t\in T$，$\sigma_{s,t}$ 为二元实值函数，$s,t\in T$，满足

(1) $\sigma_{s,t}=\sigma_{t,s}$; (1.9)

(2) 对任意的 $t_1,t_2,\cdots,t_n\in T$，矩阵

$$B_n=(\sigma_{t_i,t_j})_{1\leqslant i,j\leqslant n}\geqslant 0. \tag{1.10}$$

则必存在一个正态过程 $\{\xi_t,t\in T\}$ 使得

$$E(\xi_t)=a_t,\ \ t\in T,$$

$$E(\xi_t-a_t)(\xi_s-a_s)=\sigma_{t,s}.$$

证明 对任意的 $t_1,t_2,\cdots,t_n\in T$，记

$$B_n=(\sigma_{t_i,t_j}), \tag{1.11}$$

$$a^{(n)}=(a_{t_1},a_{t_2},\cdots,a_{t_n}),$$

则

$$f_n(u)=e^{ia^{(n)}u'-\frac{1}{2}\Sigma_k\Sigma_l\sigma_{t_k,t_l}\cdot u_k u_l}$$

$$=e^{ia^{(n)}u'-\frac{1}{2}uB_nu'} \tag{1.12}$$

是 n 维特征函数。不难看出，$\{f_n(u)\}$ 所对应的分布函数簇 $\{F_{t_1,\cdots,t_n}(x_1,\cdots,x_n)\}$ 满足K氏相容性条件：

(a) $F_{t_{\nu_1},\cdots,t_{\nu_n}}(x_{\nu_1},\cdots,x_{\nu_n})$ 对应的特征函数为

$$f_n(u_{\nu_1},u_{\nu_2},\cdots,u_{\nu_n})=e^{i\Sigma_l a_{\nu_l}u_{\nu_l}-\frac{1}{2}\Sigma_k\Sigma_l\sigma_{\nu_k,\nu_l}\cdot u_{\nu_k}\cdot u_{\nu_l}}$$

$$=e^{i\Sigma_l a_{t_l}u_{t_l}-\frac{1}{2}\Sigma_k\Sigma_l\sigma_{t_k,t_l}u_{t_k}u_{t_l}}$$

$$=f_n(u).$$

而 $f_n(u)$ 对应的分布函数为

$$F_{t_1,\cdots,t_n}(x_1,\cdots,x_n).$$

(b) $F_{t_1,\cdots,t_m}(x_1,\cdots,x_m)=F_{t_1,\cdots,t_m,t_{m+1},\cdots,t_n}(x_1,\cdots,x_m,$

$+\infty, \cdots, +\infty$)显然满足(相当于 n 维分布与 m 维($m<n$) 边缘分布的关系)。

由定理1.1的存在性定理，必存在随机过程$\{\xi_t, t\in T\}$以上述正态分布函数簇为时律，因而 $\{\xi_t, t\in T\}$ 是实值正态过程。由(1.12)知$\{\xi_t\}$以 a_t 为期望，以 $\sigma_{t,s}$ 为协方差。 ▌

上述 Gauss 过程存在定理还可以推广到复正态过程的场合。

定义1.3 设$(\xi_1, \xi_2, \cdots, \xi_n)$为$n$ 元复随机变量，称它们是遵从 n 元复 Gauss 分布的，假若它们的实虚部的联合$(\eta_1, \zeta_1, \eta_2, \zeta_2, \cdots, \eta_n, \zeta_n)$是遵从 $2n$ 元的 Gauss 分布。设$\{\xi_t, t\in T\}$ 是复值随机过程，如果对任意的正整数 n ,$(\xi_{t_1}, \xi_{t_2}, \cdots, \xi_{t_n})$是遵从复 Gauss 分布的，则称$\{\xi_t, t\in T\}$为复 Gauss 过程。

定理1.3 设 T 为足标集，a_t 为定义在 T 上的复值函数，$\sigma_{s,t}$ 为定义在 $T\times T$ 上的满足以下二个条件的二元复值函数：

(1) $\sigma_{s,t}=\bar{\sigma}_{t,s}$, $(s,t)\in T\times T$, (1.13)

(2) 对 $t_1, t_2, \cdots, t_n\in T$,

$$B_n=(\sigma_{t_i,t_j})_{1\leqslant i,j\leqslant n} \geqslant 0, \tag{1.14}$$

则必存在一复 Gauss 过程$\{\xi_t, t\in T\}$满足①

$$E(\xi_t)=a_t, \tag{1.15}$$

$$E(\xi_t-a_t)\overline{(\xi_s-a_s)}=\sigma_{t,s}. \tag{1.16}$$

此定理的证明可在[1],[2]中找到，此处不细叙。

4° 宽平稳过程与严平稳过程

以下我们将着重介绍与时间序列分析的理论及方法有密切关系的平稳随机过程的概念。

定义1.4 设$\{\xi_t, t\in T\}$为复值随机过程，其期望与协方差存在，并满足

(1) $E\xi_t\equiv a$，对一切 $t\in T$,

---

① 复值变量的期望为实虚部期望之和，即 $E\xi_t=E\eta_t+iE\zeta_t$。

(2) $E(\xi_{t+\tau}-a)\overline{(\xi_t-a)}=R_\xi(\tau)$，对 $t,t+\tau\in T$ 成立，则称$\{\xi_t,t\in T\}$为宽平稳过程，$R_\xi(\tau)$称为ξ_t的协方差函数。若不混淆亦简记为$R(\tau)$。

由定义1.4可看出：第一条要求表明过程的期望值不随时间的推移而变化；第二条要求表明过程前后两个时刻的线性相关性(统计意义)的强弱只依赖这两个时刻间的间隔而与它们所处的位置t无关(参看图1.2)。这类过程在工程技术、自然科学和医学

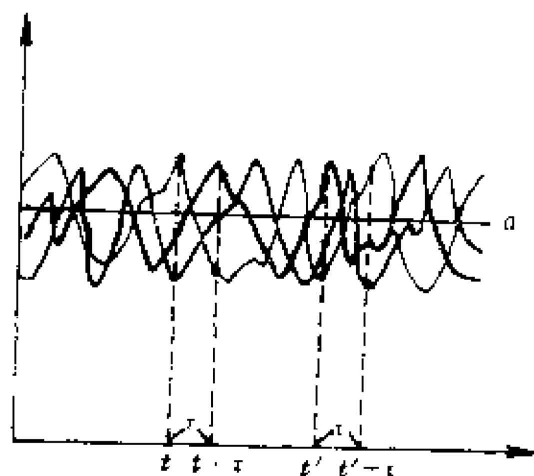


图 1.2

等领域中具有极其广泛的代表性，它反映了一个系统处于稳态工作条件下的统计性质(稳态并不是指固定不变，系统的工作仍具随机性)。

通常称

$$B(\tau)=E\xi_{t+\tau}\overline{\xi}_t \tag{1.17}$$

为宽平稳过程的相关函数，显见有

$$R(\tau)=B(\tau)-|a|^2$$

成立。当$a=0$时，相关函数与协方差函数等价。以后如不声明，我们皆指过程的期望为零。若不然，可令

$$x_t=\xi_t-a,\qquad t\in T,$$

则

$$Ex_t = 0,\ B_x(\tau) = R_x(\tau) = R_\xi(\tau).$$

以下我们举几个宽平稳过程的例子。

例1.2 白噪声序列。

设$\{\xi_t, t=0,\pm1,\pm2,\cdots\}$是独立同分布序列(实)，具有二阶矩，并满足

$$E\xi_t = 0, \quad t=0,\pm1,\pm2,\cdots,$$

$$R(\tau) = E\xi_{t+\tau}\xi_t = \begin{cases} 1, & \tau = 0; \\ 0, & \tau \neq 0. \end{cases} \tag{1.18}$$

显然，ξ_t 满足定义1.4，因而是宽平稳过程。

例1.3 随机振荡(随机相位)。

设

$$\xi_t = Ae^{i(\omega t+\theta)},\ \theta \sim U(0,2\pi), \quad t \in R_1 \tag{1.19}$$

(θ 遵从$(0,2\pi)$上的均匀分布)，$A>0$和 ω 均为常数，则$\{\xi_t, t\in \boldsymbol{R}_1\}$是宽平稳过程。

事实上可验证平稳性的两条件如下：

$$E(\xi_t) = Ae^{i\omega t}E(e^{i\theta})$$

$$= Ae^{i\omega t}\int_0^{2\pi} e^{i\theta}\ \frac{d\theta}{2\pi} \equiv 0, \quad 对一切\ t \in \boldsymbol{R}_1,$$

$$E\xi_{t+\tau}\bar{\xi}_t = E(Ae^{i[\omega(t+\tau)+\theta]}\bar{A}e^{-i(\omega t+\theta)})$$

$$= |A|^2e^{i\omega\tau} = A^2e^{i\omega\tau}, \quad t,\tau \in \boldsymbol{R}_1. \tag{1.20}$$

式(1.20)表明与 t 无关。

进一步若考虑

$$\xi_t = \sum_{k=1}^{N} A_k e^{i(\omega_k t+\theta_k)}, \quad t \in \boldsymbol{R}_1, \tag{1.21}$$

A_k, ω_k 为常数，$\{\theta_k\}$是 n 个相互独立同 $U(0,2\pi)$分布的随机变量，则 ξ_t 也是宽平稳过程。

例1.4 随机振荡(随机振幅)。

设 $\xi_t = \xi e^{i\omega t}$，$\omega$ 为常数，$E|\xi|^2 < +\infty$，$E\xi = 0$，则$\{\xi_t, t\in \boldsymbol{R}_1\}$是宽平稳过程。因为

$$E\xi_t = e^{i\omega t}E\xi = 0, \quad t\in \boldsymbol{R}_1,$$
$$E(\xi_{t+\tau}\bar{\xi}_t) = E|\xi|^2 e^{i\omega\tau}, \quad t,\tau\in \boldsymbol{R}_1$$

与 t 无关。

进一步若考虑

$$\xi_t = \sum_{k=1}^{N}\xi_k e^{i\omega_k t}, \quad t\in \boldsymbol{R}_1, \tag{1.22}$$

其中ω_k 为常数，$\{\xi_k\}$是N个相互独立、均值为零且具有二阶矩的随机变量，则$\{\xi_t, t\in \boldsymbol{R}_1\}$也是宽平稳过程。

例1.5 白噪声的过滤。

设有一个滤波器，它的滤波权系数(实)为$C_0, C_1, \cdots, C_N$。今将白噪声列$\{\varepsilon_t, t=0, \pm 1, \cdots\}$送入该滤波器，则其输出(参看图1.3)可表为

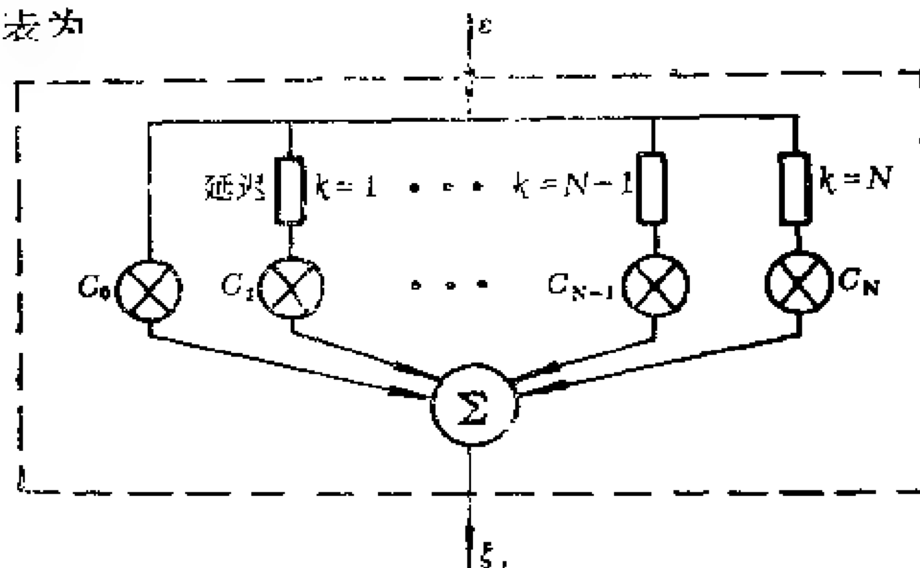


图 1.3

$$\xi_t = \sum_{k=0}^{N} C_k \varepsilon_{t-k}, \quad t=0, \pm 1, \cdots, \tag{1.23}$$

它是宽平稳列。事实上：

$$E(\xi_t) = \sum_{k=0}^{N} C_k E(\varepsilon_{t-k}) \equiv 0, \quad t=0, \pm 1, \cdots,$$

$$E(\xi_{t+\tau}\xi_t) = \sum_{k=0}^{N}\sum_{l=0}^{N} C_k C_l E(\varepsilon_{t+\tau-k}\varepsilon_{t-l}).$$

利用白噪声的(1.18)式，则

$$R_\xi(\tau)=\sum_{\substack{l=0\\(0\leqslant\tau+l\leqslant N)}}^{N}C_lC_{\tau+l}$$

$$=\begin{cases}\sum\limits_{l=\max(0,-\tau)}^{\min(N,N-\tau)}C_lC_{l+\tau}, & \text{当 } 0\leqslant|\tau|\leqslant N\\ 0, & \text{当 } |\tau|>N\end{cases}\qquad(1.24)$$

与 t 无关。

以后我们还将介绍多种宽平稳过程的模型。以下我们转入严平稳过程的介绍：

定义1.5 设$\{\xi_t,t\in T\}$是随机过程，假若对任意的$n>0$，$t_1,t_2,\cdots,t_n\in T$及τ $(t_{i+\tau}\in T,i=1,2,\cdots,n)$，$(\xi_{t_1},\xi_{t_2},\cdots,\xi_{t_n})$ 和 $(\xi_{t_1+\tau},\xi_{t_2+\tau},\cdots,\xi_{t_n+\tau})$具有相同的分布函数

$$F_{t_1,t_2,\cdots,t_n}(x_1,x_2,\cdots,x_n)=F_{t_1+\tau,t_2+\tau,\cdots,t_n+\tau}(x_1,x_2,\cdots,x_n),\qquad(1.25)$$

其中$(x_1,x_2,\cdots,x_n)\in \boldsymbol{R}_n$，则称$\{\xi_t\}$是严平稳过程。

比较定义1.4和定义1.5可以看出：宽平稳过程对时间推移的不变性(表现不依赖于 t)是表现在统计平均的一、二阶矩上，而严平稳过程对时间推移的不变性则表现在概率分布上(见(1.25)式)。显然，这两者的要求是不一样的，一般说来严平稳要求的条件比宽平稳要严得多。但是须注意的是：

1. 严平稳未必是宽平稳。因为严平稳的定义中并不断言二阶矩存在，而宽平稳的前提是协方差函数存在。

2. 宽平稳未必是严平稳。因为一、二阶矩对时间推移的不变性，一般说来并不能得到(1.25)式。例如：$\{\xi_k\}$是相互独立的一串序列，$E\xi_k\equiv0$，$\xi_1\sim U(-1,1)$，则$\sigma^2_{\xi_1}=\frac{1}{3}$；而对$i=2,3,\cdots$，$\xi_i\sim N\left(0,\frac{1}{3}\right)$，于是

$$E\xi_k \equiv 0,$$

$$E\xi_i\xi_j = \begin{cases} 0, & i \neq j; \\ \dfrac{1}{3}, & i = j. \end{cases}$$

故$\{\xi_k\}$是宽平稳列。但取$\tau = 1$，分布函数

$$F_{1,2}(x,y) \neq F_{2,3}(x,y).$$

3. 严平稳过程若具有二阶矩，则必为宽平稳过程。因为

$$E\xi_t = \int_{R_1} x\mathrm{d}F_t(x) = \int_{R_1} x\mathrm{d}F_{t+\tau}(x) \equiv a,$$

$$E\xi_{t+\tau}\xi_t = \iint_{R_2} x_1x_2\mathrm{d}F_{t+\tau,t}(x_1,x_2)$$

$$= \iint_{R_2} x_1x_2\mathrm{d}F_{\tau,0}(x_1,x_2) = B(\tau)$$

与t无关。

4. 对正态过程，宽平稳与严平稳等价。由于正态过程是有二阶矩的，因此，若$\{\xi_t, t\in T\}$是严平稳，则由3知必为宽平稳。反之，若正态过程是宽平稳，则

$$E\xi_t \equiv a,\quad t\in T,$$

$$E(\xi_{t_i} - a)(\xi_{t_j} - a) = R(t_i - t_j) = R((t_i+\tau) - (t_j+\tau))$$

$$= E(\xi_{t_i+\tau} - a)(\xi_{t_j+\tau} - a)$$

$$t_i \in T, i = 1,2,\cdots,n.$$

表明$(\xi_{t_1}, \xi_{t_2},\cdots,\xi_{t_n})$和$(\xi_{t_1+\tau}, \xi_{t_2+\tau},\cdots,\xi_{t_n+\tau})$具有相同的协方差矩阵和均值向量，因而

$$F_{t_1,\cdots,t_n}(x_1,\cdots,x_n) = F_{t_1+\tau,\cdots,t_n+\tau}(x_1,\cdots,x_n).$$

$\{\xi_t, t\in T\}$是严平稳过程。

今后主要研究对象是宽平稳过程，并简称为平稳过程，如不声明均假定是零均值的。

5° 其它常见的随机过程

1. 马氏链

先看一个简单模型，设有四个罐，编号为1,2,3,4，每罐皆装有

印号为1,2,3,4的小球，第 i 号罐的内容是 1 号球有 $\nu_1^{(i)}$ 个，2 号球有 $\nu_2^{(i)}$ 个，…，4 号球有 $\nu_4^{(i)}$ 个。开始时任选一罐任取一球，设取中第 j 号球($1\leqslant j\leqslant 4$)，下一步则转移到第 j 号罐去取球(每次看后放回)。若用 ξ_l 表示第 l 次取中球号的随机变量，显然有以下条件概率等式成立

$$P\{\xi_{m+1}=i_{m+1}|\xi_m=i_m,\xi_{m-1}=i_{m-1},\cdots,\xi_1=i_1\}$$
$$=P\{\xi_{m+1}=i_{m+1}|\xi_m=i_m\} \quad (1\leqslant i_l\leqslant 4). \tag{1.26}$$

(1.26)式称为马氏性。上述模型是马氏链的简单模型，一般地我们有以下定义

定义1.6 设$\{\xi_n,n=0,1,2,\cdots\}$为定义在某概率空间上的随机列，ξ_n 取非负整数值，如果对任意的正整数 l,m,k 以及任意的非负整数

$$0\leqslant j_1<j_2<\cdots<j_l<m \tag{1.27}$$

皆有

$$P\{\xi_{m+k}=i_{m+k}|\xi_m=i_m,\xi_{j_l}=i_{j_l},\cdots,\xi_{j_1}=i_{j_1}\}$$
$$=P\{\xi_{m+k}=i_{m+k}|\xi_m=i_m\} \tag{1.28}$$

对任意的非负整数 $i_{m+k},i_m,i_{j_l},\cdots,i_{j_1}$ 成立①，则称$\{\xi_n\}$为马氏链。

由定义可看出马氏链的如下无后效性：如果已知现在(第 m 个时刻)的实验结果为 i_m，则将来(第 $m+k$ 时刻)的结果 i_{m+k} 只与现在的结果i_m 有关而不依赖于过去$\{j_k\}$诸时刻的结果($i_{j_1},i_{j_2},\cdots,i_{j_l}$)。

2. 独立增量过程

设$\{\xi_t,t\in T\}$是随机过程，如果对 $t_1<t_2<\cdots<t_n$ ($t_i\in T$)，下列随机变量之差

$$\xi_{t_2}-\xi_{t_1},\xi_{t_3}-\xi_{t_2},\cdots,\xi_{t_n}-\xi_{t_{n-1}}$$

---

① 若 ξ 取值为部分非负整数，则 $\{i_n\}$ 亦应在可能取值范围内。

是相互独立的，则称$\{\xi_t\}$为独立增量过程。以下二个过程是常见的独立增量过程。

Poisson 过程：$\xi_t, t\in T=[0,+\infty)$是随机过程，如果它满足以下条件就称为是 Poisson 过程：

(a) 是取非负整数值的独立增量过程；

(b) 对任意的① $0\leqslant s<t$，$\xi_t-\xi_s$遵从 Poisson 分布，即

$$P\{\xi_t-\xi_s=k\}=\frac{(\lambda(t-s))^k}{k!}e^{-\lambda(t-s)},\quad k=0,1,2,\cdots,\quad (1.29)$$

其中$\lambda>0$，$\lambda(t-s)$为数学期望。

Wiener 过程：称随机过程$\{\xi_t, t\in[0,+\infty)\}$为 Wiener 过程，假若它满足

(a) 是独立增量过程；

(b) 对任意的$0\leqslant s<t$，$\xi_t-\xi_s$是均值为零，方差为$\sigma^2(t-s)$的正态随机变量，即

$$\xi_t-\xi_s\sim N(0,\sigma^2\cdot(t-s)),\sigma>0. \quad (1.30)$$

随着科学技术的飞速发展，与之有密切联系的随机过程也在日新月异地迅速发展，读者在其它书籍和文献中还将遇到许多类型的随机过程，我们这里不可能一一介绍，读者可参看[1],[2],[3],[4]。

§2 平稳随机过程的相关函数

1° 具有有限方差的随机变量所组成的 Hilbert 空间

在泛函分析[5]或实变函数[6]中介绍过 Hilbert 空间的知识。简言之，设H是一个集合，它是线性的(元素间可进行线性运算)，具有内积定义的(满足内积四条件)并且是完备的(Cauchy 列皆有属于H的极限点)空间，称H为 Hilbert 空间。最典型的 Hilbert

---

① 通常由$t=0$开始的 Poisson，Wiener 过程均假定$\xi_0=0$。

空间有

$L^2(\mathrm{d}\lambda)$：定义在有限区间(a,b)上的平方可积函数的全体，在内积

$$(f,g)=\int_a^b f\bar{g}\,\mathrm{d}\lambda,\qquad f,g\in L^2(\mathrm{d}\lambda) \tag{1.31}$$

定义下组成了一个 Hilbert 空间。

l^2：全体复数列$(a_1,a_2,\cdots)=\boldsymbol{\alpha}$，平方可和$\sum\limits_i|a_i|^2<+\infty$，在内积

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\sum_j a_j\bar{b}_j,\quad \boldsymbol{\alpha},\boldsymbol{\beta}\in l^2 \tag{1.32}$$

定义下是 Hilbert 空间。

在随机过程论中，经常要运用泛函分析的知识来解决概率论中的问题。以下介绍具有二阶矩的随机变量所组成的Hilbert空间。

设定义在概率空间$(\Omega,\mathscr{F},P)$上的具有二阶矩的随机变量的全体记为H。对$\xi,\eta\in H$，定义内积为

$$(\xi,\eta)=E(\xi\bar{\eta}),\qquad \xi,\eta\in H, \tag{1.33}$$

则H是 Hilbert 空间。以下逐条验证各条件：

(1) 线性性质：设$\xi_1,\xi_2\in H$，α,β为常数，则由概率论的知识可证以下的 Minkovski 不等式成立

$$\begin{aligned}E|\eta_1+\eta_2|^2&=|E(\eta_1+\eta_2)\overline{(\eta_1+\eta_2)}|\\&=|E|\eta_1|^2+E|\eta_2|^2+E(\eta_1\bar{\eta}_2)+E(\eta_2\bar{\eta}_1)|\\&\leqslant E|\eta_1|^2+E|\eta_2|^2+2|E(\eta_1\bar{\eta}_2)|\\&\leqslant E|\eta_1|^2+E|\eta_2|^2+2\sqrt{E|\eta_1|^2E|\eta_2|^2}\ ^{①}\\&=(\sqrt{E|\eta_1|^2}+\sqrt{E|\eta_2|^2})^2.\end{aligned} \tag{1.34}$$

所以

$$E|\alpha\xi_1+\beta\xi_2|^2\leqslant(|\alpha|\sqrt{E|\xi_1|^2}+|\beta|\sqrt{E|\xi_2|^2})^2<+\infty. \tag{1.35}$$

表明$\alpha\xi_1+\beta\xi_2\in H$。

---

① 由$|\rho_{\eta_1\eta_2}|=|E(\eta_1\bar{\eta}_2)|/\sqrt{E|\eta_1|^2\cdot E|\eta_2|^2}\leqslant 1$或直接验证。

(2) 满足内积的四条性质：

$$(\xi_1+\xi_2,\eta)=E(\xi_1+\xi_2)\bar{\eta}=E\xi_1\bar{\eta}+E\xi_2\bar{\eta}$$
$$=(\xi_1,\eta)+(\xi_2,\eta),$$
$$(a\xi,\eta)=E(a\xi\bar{\eta})=aE(\xi\bar{\eta})=a(\xi,\eta).$$
$$(\xi,\eta)=E\xi\bar{\eta}=\overline{E(\bar{\xi}\eta)}=\overline{E(\eta\bar{\xi})}=\overline{(\eta,\xi)},$$
$$(\xi,\xi)\geqslant 0.$$

$E|\xi|^2=(\xi,\xi)=0$ 表明概率为 1 ξ 取 0 值（实虚部分别取 0 值）。在H中的 0 元素皆指在概率为 1 的意义下，由此，H中两个元素相等也是在概率为 1 的意义下，故

$$(\xi,\xi)=0\Longleftrightarrow\xi=0.$$

(3) 完备性：设 $\xi_n\in H$，它是 Cauchy 列，即

$$E|\xi_n-\xi_m|^2\to 0,\quad n,m\to\infty, \tag{1.36}$$

今证[①]必在H中存在一元素 ξ 使

$$E|\xi_n-\xi|^2\to 0,\quad n\to\infty. \tag{1.37}$$

首先，由车贝谢夫不等式，对 $\sigma>0$ 有

$$P\{|\xi_n-\xi_m|\geqslant\sigma\}\leqslant\frac{1}{\sigma^2}E|\xi_n-\xi_m|^2\to 0\ (n,m\to\infty),$$

因而必存在子序列 $\xi_{k_n}\xrightarrow{a.s}\xi\ (n\to\infty)$(见[7])，因而对固定的 m，

$$\xi_m-\xi_{k_n}\xrightarrow{a.s}\xi_m-\xi\quad(n\to\infty). \tag{1.38}$$

但

$$E|\xi_m-\xi_{k_n}|^2\to 0\quad(m,n\to\infty), \tag{1.39}$$

由 Fatou 引理和(1.38)，(1.39)式可得：

$$E|\xi_m-\xi|^2=E(\liminf_n|\xi_m-\xi_{k_n}|^2)$$
$$\leqslant\liminf_n E|\xi_m-\xi_{k_n}|^2<\delta\quad(\text{当} m \text{充分大}).$$

① 以下用到的概率论知识可参看[2]的附录。

由此，存在$m_0>0$使得当$m\geqslant m_0$以后$E|\xi_m-\xi|^2$都有穷，即$\xi-\xi_m\in H$。又由$\xi_m\in H$得

$$\xi=\xi_m+(\xi-\xi_m)\in H.$$

在H内记

$$\|\xi\|=\sqrt{E|\xi|^2},\qquad \xi\in H,\tag{1.40}$$

则(1.34)的 Minkovski 不等式即变成大家熟知的三角不等式

$$\|\eta_1+\eta_2\|\leqslant\|\eta_1\|+\|\eta_2\|.\tag{1.41}$$

在此范数意义下H中的收敛就是均方意义的，以后用(L^2)示之。

有时嫌H太大，对一个具有二阶矩的过程$\{\xi_t,t\in T\}$而言，可以研究更具体的 Hilbert 空间，这就是由ξ_t在(L^2)意义下张成的线性闭包，其过程如下：

$$\{\xi_t,t\in T\}\overset{\text{扩}}{\Longrightarrow}L[\xi_t]=\left\{\sum_{k=1}^{n}c_k^{(n)}\xi_{t_k^{(n)}}\,\middle|\,n>0,t_k^{(n)}\in T,c_k^{(n)}\text{常数}\right\}$$

$$\overset{\text{扩}}{\Longrightarrow}(L[\xi_t])^c\triangleq\mathscr{L}\{\xi_t,t\in T\}.\tag{1.42}$$

$(\quad)^c$是指对线性集$L[\xi_t]$求闭包，即将其一切可能的序列的极限点(由H是完备的)补充进去。显见$\mathscr{L}\{\xi_t,t\in T\}\triangleq H_\xi\subset H$并且是一个子 Hilbert 空间。有时还嫌$H_\xi$太大，需考虑更小的如下的子空间：

$$H_\xi(s)=\mathscr{L}\{\xi_t,t\in s\},\quad s\subset T.\tag{1.43}$$

s的元素可以是无穷多个也可以是有限个，后者的场合

$$\mathscr{L}\{\xi_t,t\in s\}=L[\xi_t,t\in s].$$

2° 平稳过程相关函数的性质

定理1.4 设$\{\xi_t,t\in T\}$为平稳过程($E\xi_t\equiv0$)，其相关函数$R_\xi(\tau)$具有以下性质

(1) $R_\xi(0)\geqslant0$; (1.44)

(2) $|R_\xi(\tau)|\leqslant R_\xi(0)$; (1.45)

(3) $R_\xi(-\tau)=\overline{R_\xi(\tau)}$; (1.46)

(4) 对 $t_1,t_2,\cdots,t_n\in T$，矩阵

$$(R_\xi(t_i-t_j))_{1\leqslant i,j\leqslant n}\geqslant 0. \tag{1.47}$$

证明 (1) 由定义 $R_\xi(0)=E|\xi_t|^2\geqslant 0$。

(2)
$$\begin{aligned}|R_\xi(\tau)| &= |E\xi_{t+\tau}\bar{\xi}_t|\\ &= |(\xi_{t+\tau},\xi_t)|\\ &\leqslant \|\xi_{t+\tau}\|\|\xi_t\| \quad (\text{由 Schwartz 不等式})\\ &= \sqrt{E|\xi_0|^2}\sqrt{E|\xi_0|^2}=R_\xi(0)\quad(\text{利用平稳性}).\end{aligned}$$

(3) $\overline{R_\xi(\tau)}=\overline{(\xi_{t+\tau},\xi_t)}=(\xi_t,\xi_{t+\tau})=R_\xi(-\tau)$。

(4) 任取复数 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)$，二次型

$$\begin{aligned}\boldsymbol{\alpha}(R_\xi(t_i-t_j))\boldsymbol{\alpha}' &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_i\bar{a}_jR_\xi(t_i-t_j)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}a_i(\xi_{t_i},\xi_{t_j})\bar{a}_j\\ &= \left(\sum_{i=1}^{n}a_i\xi_{t_i},\sum_{j=1}^{n}a_j\xi_{t_j}\right)\\ &= \left\|\sum_{i=1}^{n}a_i\xi_{t_i}\right\|^2\geqslant 0.\quad \blacksquare\end{aligned}$$

在随机过程的研究中，我们经常要遇到各种与“连续性”有关的问题，其中常用的一种是均方连续。

定义1.7 设$\{\xi_t,t\in T\}$，$T\subset \boldsymbol{R}_1$是具有连续足标参数的平稳过程，若对 $t_0\in T$，有

$$\lim_{h\to 0}\|\xi_{t_0+h}-\xi_{t_0}\|=0, \tag{1.48}$$

就称$\{\xi_t\}$在 t_0 点均方连续，如果对每个 $t\in T$ 都均方连续，就称$\{\xi_t\}$在 T 上是均方连续的。

如何判断一个过程是否是均方连续的？以下定理给出了这类判别的依据。

定理1.5 设$\{\xi_t,t\in \boldsymbol{R}_1\}$是平稳过程，则以下命题等价：

(a) $\{\xi_t\}$在 $\boldsymbol{R}_1$ 上均方连续；

(b) $\{\xi_t\}$在 $t=0$ 均方连续；

(c) $R_\xi(\tau)$在 $\boldsymbol{R}_1$ 上连续；

(d) $R_\xi(\tau)$在 $\tau=0$ 连续。

证明 显见(a)→(b)自然成立。

(b)→(c)可由以下推导得出($E\xi_t\equiv 0$)：

$$\begin{aligned}|R_\xi(t+\tau)-R_\xi(t)| &= |(\xi_{t+\tau},\xi_0)-(\xi_t,\xi_0)|\\ &= |(\xi_{t+\tau}-\xi_t,\xi_0)|\\ &\leqslant \|\xi_{t+\tau}-\xi_t\|\|\xi_0\|,\end{aligned}\tag{1.49}$$

然而

$$\begin{aligned}&\|\xi_{t+\tau}-\xi_t\|^2\\ &=(\xi_{t+\tau}-\xi_t,\xi_{t+\tau}-\xi_t)\\ &=(\xi_{t+\tau},\xi_{t+\tau})-(\xi_t,\xi_{t+\tau})-(\xi_{t+\tau},\xi_t)+(\xi_t,\xi_t)\\ &=(\xi_\tau,\xi_\tau)-(\xi_0,\xi_\tau)-(\xi_\tau,\xi_0)+(\xi_0,\xi_0)\\ &=(\xi_\tau-\xi_0,\xi_\tau-\xi_0)=\|\xi_\tau-\xi_0\|^2,\end{aligned}\tag{1.50}$$

将(1.50)代入(1.49)得

$$|R_\xi(t+\tau)-R_\xi(t)|\leqslant\|\xi_\tau-\xi_0\|\cdot\|\xi_0\|\to 0\quad(\tau\to 0).\tag{1.51}$$

(1.51)右端的断言是由条件(b)得到的，由于t的任意性，(c)成立。

(c)→(d)是显然的。

(d)→(a)。由(1.50)可得

$$\begin{aligned}\|\xi_{t-\tau}-\xi_t\|^2&=\|\xi_\tau-\xi_0\|^2\\ &=2R_\xi(0)-R_\xi(\tau)-R_\xi(-\tau)\to 0\quad(\text{当 }\tau\to 0).\end{aligned}$$

由此表明对任意的 $t\in\boldsymbol{R}_1$，ξ_t均方连续。 ▎

须注意的是：ξ_t在 $\boldsymbol{R}_1$上(L^2)连续并不能断言它的观测现实x_t(样本)函数是连续的。

例1.6 随机电报信号(见[17])。

对任一时刻 $t\in\boldsymbol{R}_1$，随机过程 $\xi(t)$只取两种状态“1”和“0”，机会均等，即

$$\xi(t) \sim \begin{pmatrix} 1, & 0 \\ \frac{1}{2}, & \frac{1}{2} \end{pmatrix}, \quad t \in \boldsymbol{R}_1, \qquad (1.52)$$

在T时间间隔内，波形(状态)可能发生变化的次数μ_T遵从以下的 Poisson 分布

$$P\{\mu_T = k\} = \frac{(\lambda T)^k}{k!} e^{-T\lambda}, \quad k = 0, 1, 2, \cdots. \qquad (1.53)$$

其中T, λ均大于0，$\xi(t)$与μ_T是独立的(即任意T间隔内波形变化的次数不受t时刻过程所处状态的影响)。试证$\xi(t)$是均方连续的平稳过程(参看图1.4)。

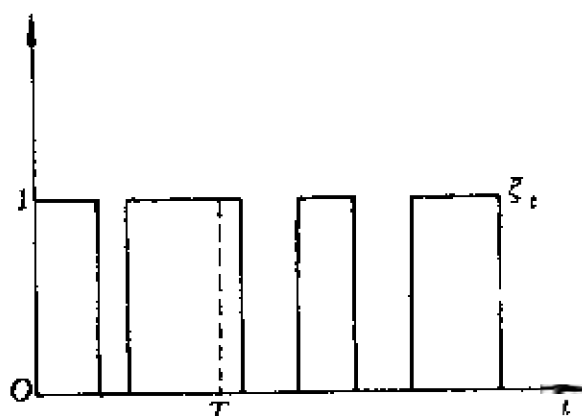


图 1.4

证明 先证明$\xi(t)$是平稳过程。

1. $E\xi(t) = \frac{1}{2}$，一切$t \in R_1$。

2. $E\xi(t+\tau)\xi(t) = P\{\xi(t) = 1, \ \xi(t+\tau) = 1\}$

$= P\{\xi(t) = 1, \text{ 在}|\tau|\text{间隔内 } \mu_{|\tau|} = \text{偶数}\}$

$= P\{\xi(t) = 1\} P\{\mu_{|\tau|} = \text{偶数}\}$

$$= \frac{1}{2} \sum_{\substack{k=0 \\ (k\text{偶})}}^{\infty} \frac{(\lambda|\tau|)^k}{k!} e^{-\lambda|\tau|}$$

$$= \frac{1}{2} \left\{ \sum_{k=0}^{\infty} \left[\frac{(\lambda|\tau|)^k + (-\lambda|\tau|)^k}{2k!} \right] e^{-\lambda|\tau|} \right\}$$

$$= \frac{1}{4}\{e^{-\lambda|\tau|}e^{\lambda|\tau|} + e^{-\lambda|\tau|}e^{-\lambda|\tau|}\}$$

$$= \frac{1}{4}(1 + e^{-2\lambda|\tau|}) = B_\xi(\tau).$$

故

$$R_\xi(\tau) = B_\xi(\tau) - (E\xi(t))^2 = \frac{1}{4}(1 + e^{-2\lambda|\tau|}) - \left(\frac{1}{2}\right)^2$$

$$= \frac{1}{4}e^{-2\lambda|\tau|}, \quad \text{一切 } \tau\text{(见图1.5)}. \tag{1.54}$$

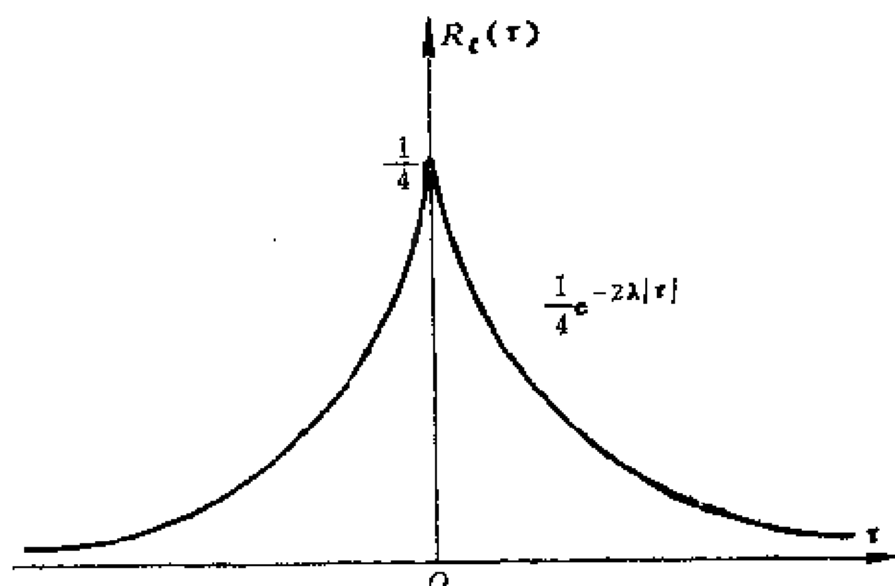


图 1.5

由 $R_\xi(\tau)$ 与 t 无关，故 $\xi(t)$ 平稳。又由 $R_\xi(\tau)$ 在 $\tau = 0$ 是连续的，故由定理1.5知，$\xi(t)$ 是均方连续的。▌

$\xi(t)$ 虽然是 (L^2) 连续的，但由 $\xi(t)$ 的定义知（图 1.4），$\xi(t)$ 的样本函数是间断函数。

3° 平稳过程相关函数的谱表示

这一节里我们将证明平稳过程的相关函数必有谱表示。这一定理有时被称为 Хинчин–Bochner 定理，我们主要针对序列场合

来证明。

定理1.6 设$\{R_\xi(n),\ n=0,\pm1,\cdots\}$是平稳序列$\{\xi_n, n=0,\pm1,\cdots\}$的相关函数($E\xi_n\equiv0$)，则$R_\xi(n)$可表为

$$R_\xi(n)=\int_{-\pi}^{\pi}e^{inu}dF_\xi(u),\quad n=0,\pm1,\cdots \tag{1.55}$$

的形式，其中dF_ξ是$[-\pi,\pi]$上的有限测度，它在$[-\pi,\pi]$的Borel集系[1]上还是唯一的。

证明 首先引入以下的二元函数

$$f(u,\rho)=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}e^{-i(n-m)u}R_\xi(n-m)\rho^{n+m} \tag{1.56}$$

$$(0<\rho<1,\quad -\pi\leqslant u\leqslant\pi).$$

易见，当取ρ固定时，$f(u,\rho)$对u是一致收敛的

$$\sum_{n=0}^{N}\sum_{m=0}^{N}|e^{-i(n-m)u}R_\xi(n-m)\rho^{n+m}|$$

$$\leqslant R_\xi(0)\left[\sum_{n=0}^{N}\rho^n\right]^2\leqslant\frac{R_\xi(0)}{(1-\rho)^2}(0<\rho<1).$$

又利用$R_\xi(n)$的非负定性（见(1.47)式），知

$$f_N(u,\rho)=\sum_{n=0}^{N}\sum_{m=0}^{N}(e^{-inu}\rho^n)R_\xi(n-m)\overline{(e^{-imu}\rho^m)}\geqslant0.$$

从而

$$\lim_{N\to\infty}f_N(u,\rho)=f(u,\rho)\geqslant0. \tag{1.57}$$

令$k=n-m$或$n=k+m$，对(1.56)作足标替换并交换求和次序可得

$$f(u,\rho)=\sum_{m=0}^{\infty}\sum_{k=-m}^{\infty}e^{-iku}R_\xi(k)\rho^{k+2m}$$

$$=\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}e^{-iku}R_\xi(k)\rho^{k+2m}$$

---

[1] 以后也用$\mathscr{B}[-\pi,\pi]$来代表。

$$+\sum_{k=-\infty}^{-1}\sum_{m=-k}^{\infty}e^{-iku}R_\xi(k)\rho^{k+2m}$$

$$=\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}e^{-iku}R_\xi(k)\rho^{k+2m}$$

$$+\sum_{k=-\infty}^{-1}\sum_{j=0}^{\infty}e^{-iku}R_\xi(k)\rho^{k+2(j-k)}$$

$$=\sum_{k=-\infty}^{\infty}e^{-iku}R_\xi(k)\left(\sum_{j=0}^{\infty}\rho^{|k|+2j}\right)$$

$$=\sum_{k=-\infty}^{\infty}e^{-iku}R_\xi(k)\frac{\rho^{|k|}}{1-\rho^2}\quad(0<\rho<1). \tag{1.58}$$

上述级数对 u 是一致收敛的，因而固定 ρ 时，$f(u,\rho)\in L_1(d\lambda)$。由(1.58)

$$R_\xi(n)\frac{\rho^{|n|}}{1-\rho^2},\qquad n=0,\pm1,\cdots \tag{1.59}$$

是 $f(u,\rho)$的 Fourier 系数(见[8])，即有

$$\frac{\rho^{|n|}}{1-\rho^2}R_\xi(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{inu}f(u,\rho)du \tag{1.60}$$

成立，或

$$\rho^{|n|}R_\xi(n)=\int_{-\pi}^{\pi}e^{inu}\,d\left(\frac{1-\rho^2}{2\pi}\int_{-\pi}^{u}f(x,\rho)dx\right). \tag{1.61}$$

令

$$F_\rho(x)=\frac{1-\rho^2}{2\pi}\int_{-\pi}^{x}f(u,\rho)du\geqslant0. \tag{1.62}$$

它是$[-\pi,\pi]$上的连续函数(对 x)并且是非降的，

$$\begin{cases}F_\rho(-\pi)=0,\\ F_\rho(\pi)=\dfrac{1-\rho^2}{2\pi}\displaystyle\int_{-\pi}^{\pi}f(u,\rho)du=R_\xi(0).\end{cases}$$

于是可将 $F_\rho(x)$规范为分布函数

$$\bar{F}_\rho(x)=\begin{cases}F_\rho(x)/R_\xi(0), & -\pi\leqslant x\leqslant\pi;\\ 1, & \pi<x;\\ 0, & x<-\pi.\end{cases} \tag{1.63}$$

由

$$|\bar{F}_\rho(x)|\leqslant 1,\quad 0<\rho<1$$

知，$\{\bar{F}_\rho(x)\}$是一致有界的分布函数簇. 取一串 $\rho_n\nearrow 1$，则由 Helly 定理（见[9]）知，分布函数簇$\{\bar{F}_{\rho_n}(x)\}$必存在一串弱收敛的子序列$\{\bar{F}_{\rho_{n_k}}(x)\}$收敛到某个分布函数 $\bar{F}_\xi(x)$：

$$\bar{F}_{\rho_{n_k}}(x)\xrightarrow{W}\bar{F}_\xi(x),\qquad x\in\boldsymbol{R}_1. \tag{1.64}$$

又由特征函数的定理（见[9]）知，必有

$$\lim_{k\to\infty}\int_{-\pi}^{\pi}e^{ixt}\,d\bar{F}_{\rho_{n_k}}(x)=\int_{-\pi}^{\pi}e^{ixt}d\bar{F}_\xi(x).$$

两边乘以 $R_\xi(0)$，得

$$\lim_{k\to\infty}\int_{-\pi}^{\pi}e^{ixt}\,dF_{\rho_{n_k}}(x)=\int_{-\pi}^{\pi}e^{ixt}dF_\xi(x). \tag{1.65}$$

再由 $\rho_{n_k}\nearrow 1$及(1.61)知

$$R_\xi(n)=\lim_{k\to\infty}\rho_{n_k}^{|n|}R_\xi(n)=\lim_{k\to\infty}\int_{-\pi}^{\pi}e^{ixn}\,dF_{\rho_{n_k}}(x)$$

$$=\int_{-\pi}^{\pi}e^{inx}dF_\xi(x). \tag{1.66}$$

以下再证明 dF_ξ 测度的唯一性问题.

假设有两个测度 dF_1,dF_2定义于$[-\pi,\pi]$上的 Borel 集系上，并有

$$R_\xi(n)=\int_{-\pi}^{\pi}e^{in\lambda}dF_1(\lambda)=\int_{-\pi}^{\pi}e^{in\lambda}dF_2(\lambda). \tag{1.67}$$

考虑定义在$[-\pi,\pi]$上的函数集合 K：

$$K=\left\{f \mid f \text{ 是Borel 可测函数并且} \int_{-\pi}^{\pi} f\mathrm{d}F_1=\int_{-\pi}^{\pi} f\mathrm{d}F_2\right\}. \tag{1.68}$$

于是

1．$\{e^{it\lambda}\}\in K$，表明K非空(等式(1.67))。

2．若$f,g\in K$，则$\alpha f+\beta g\in K$。由积分的线性性质知K是线性集。

3．若$f_n\in K$，$f_n\xrightarrow{u.c}f$(一致收敛)，则$f\in K$(因为一致收敛便可积分号下取极限)。

4．连续函数类$C\subset K$。设$f(x)$为$[-\pi,\pi]$上的连续函数，由 Wierstrass 定理(见[10])，必存在三角多项式$T_n(x)\xrightarrow{u.c}f(x)$。而$T_n(x)$只是$\{e^{ikx}\}$的线性组合，由3．知

$$T_n(x)\in K\Longrightarrow f(x)\in K\Longrightarrow C\subset K.$$

5．对任一 Borel 集$A\in\mathscr{B}[-\pi,\pi]$，考查$A$的示性函数$\chi_A(x)$。显然

$$|\chi_A(x)|\leqslant 1,\qquad x\in[-\pi,\pi].$$

由实变函数(见[6])的知识可知：必存在一串有界连续函数$\psi_n(x)$来逼近$\chi_A(x)$。因而

$$\lim_{n}\int_{-\pi}^{\pi}\psi_n(x)\mathrm{d}F_1=\int_{-\pi}^{\pi}\chi_A(x)\mathrm{d}F_1=F_1(A).$$

又

$$F_1(A)=\lim_{n}\int_{-\pi}^{\pi}\psi_n(x)\mathrm{d}F_2=\int_{-\pi}^{\pi}\chi_A(x)\mathrm{d}F_2=F_2(A),$$

由此完成了本定理的全部证明。▎

由定理1.6知：设$\{\xi_t,t=0,\pm1,\cdots\}$为平稳列，则必存在唯一的谱测度$\mathrm{d}F_\xi$与之对应。若取$A=[-\pi,\lambda],\lambda\leqslant\pi$，则

$$F_\xi(\lambda)=\int_A\mathrm{d}F_\xi(\lambda)-F_\xi\{[-\pi,\lambda)\} \tag{1.69}$$

为$\{\xi_t\}$的谱函数(非降，左连续)。如果有$\mathrm{d}F_\xi\ll\mathrm{d}\lambda$(即对 Lebesgue

测度绝对连续)，即存在非负可积函数 $f_\xi(\lambda)$ 使得

$$F_\xi\{A\}=\int_A f_\xi(\lambda)\mathrm{d}\lambda,\quad A\in\mathscr{B}[-\pi,\pi](\text{Borel集系}), \tag{1.70}$$

则以后称 $f_\xi(\lambda)$ 为过程 $\{\xi_t\}$ 的谱密度函数。此时有

$$R_\xi(n)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}n\lambda}\mathrm{d}F_\xi(\lambda)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}n\lambda}f_\xi(\lambda)\mathrm{d}\lambda,\quad n=0,\pm1,\cdots \tag{1.71}$$

成立。

什么条件能保证平稳序列有谱密度存在？一个常用的充分条件是：如果 $\{\xi_t\}$ 的相关函数满足

$$\sum_{k=-\infty}^{\infty}|R_\xi(k)|<+\infty. \tag{1.72}$$

事实上，(1.72)保证了级数

$$\sum_{k=-\infty}^{\infty}R_\xi(k)\mathrm{e}^{-\mathrm{i}k\lambda},\quad -\pi\leqslant\lambda\leqslant\pi$$

是一致(绝对)收敛的。于是可令

$$f(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}R_\xi(k)\mathrm{e}^{-\mathrm{i}k\lambda},\quad -\pi\leqslant\lambda\leqslant\pi, \tag{1.73}$$

则 $f(\lambda)\in L_1(\mathrm{d}\lambda)$。从而(1.73)是 $f(\lambda)$ 的 Fourier 级数，因此

$$R_\xi(k)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}k\lambda}f(\lambda)\mathrm{d}\lambda.$$

这表明

$$F_\xi(\lambda)=\int_{-\pi}^{\lambda}f(x)\mathrm{d}x,\quad \mathrm{d}F_\xi\ll\mathrm{d}\lambda,$$

并且谱密度 $f_\xi(\lambda)$ 就是(1.73)所定义。显见

$$f_\xi(\lambda)\geqslant0.$$

设 $\{\xi_t,t\in\boldsymbol{R}_1\}$ 是 (L^2) 连续的平稳过程，则也有类似于定理1.6的结果(见[3])，即其相关函数 $R_\xi(\tau)$ 可表成

$$R_\xi(\tau)=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}\tau\lambda}\mathrm{d}F_\xi(\lambda),\quad -\infty<\tau<+\infty. \tag{1.74}$$

其中测度 $\mathrm{d}F_\xi$ 对 $\boldsymbol{R}_1$ 上的 Borel 集系 $\mathscr{B}_1$ 还是唯一的。如果 ξ_t 的相关函数还是绝对可积的

$$\int_{-\infty}^{\infty}|R_\xi(\tau)|\,\mathrm{d}\tau<+\infty, \tag{1.75}$$

则 $\mathrm{d}F_\xi \ll \mathrm{d}\lambda$，存在非负可积函数 $f_\xi(\lambda)$ 使

$$F_\xi(A)=\int_A f_\xi(\lambda)\mathrm{d}\lambda, \qquad A\in\mathscr{B}_1. \tag{1.76}$$

通常取

$$F_\xi(\lambda)=F_\xi\{x<\lambda\}, \qquad -\infty<x,\lambda<+\infty \tag{1.77}$$

为 ξ_t 的谱函数，于是 $f_\xi(\lambda)$ 就称为 ξ_t 的谱密度。综合以上结果有：

定理1.7 (Wiener-Хинчин) 设 $\{\xi_n, n=0,\pm1,\cdots\}$ 是平稳列，其相关函数满足(1.72)，则 ξ_n 必有非负谱密度函数 $f_\xi(\lambda)$，并且 $R_\xi(\tau)$ 和 $f_\xi(\lambda)$ 是 Fourier 变换的关系：

$$\begin{cases} R_\xi(k)=\displaystyle\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}k\lambda}f_\xi(\lambda)\mathrm{d}\lambda, & k=0,\pm1,\cdots, \\ f_\xi(\lambda)=\dfrac{1}{2\pi}\displaystyle\sum_{k=-\infty}^{\infty}R_\xi(k)\mathrm{e}^{-\mathrm{i}k\lambda}, & -\pi\leqslant\lambda\leqslant\pi. \end{cases} \tag{1.78}$$

推论 如果 $\{\xi_n, n=0,\pm1,\cdots\}$ 是实平稳列，$R_\xi(\tau)$ 绝对可和，则谱密度 $f_\xi(\lambda)$ 必存在，并

(a) $f_\xi(\lambda)=f_\xi(-\lambda), \quad -\pi\leqslant\lambda\leqslant\pi,$ (1.79)

(b)
$$\begin{cases} f_\xi(\lambda)=\dfrac{1}{2\pi}\left\{R_\xi(0)+2\displaystyle\sum_{k=1}^{\infty}R_\xi(k)\cos k\lambda\right\} \\ R_\xi(n)=2\displaystyle\int_0^{\pi}f_\xi(\lambda)\cos n\lambda\,\mathrm{d}\lambda \end{cases} \tag{1.80}$$

定理1.8 (Wiener-Хинчин) 设 $\{\xi_t, t\in\boldsymbol{R}_1\}$ 是 (L^2) 连续的平稳过程，其相关函数 $R_\xi(\tau)$ 满足(1.75)，则 ξ_t 必有非负的谱密度函数 $f_\xi(\lambda)$，并 $R_\xi(\tau)$ 和 $f_\xi(\lambda)$ 是一对 Fourier 变换关系：

$$\begin{cases} R_\xi(\tau)=\int_{-\infty}^{\infty} e^{i\tau\lambda} f_\xi(\lambda) d\lambda, & \tau\in \boldsymbol{R}_1 \\ f_\xi(\lambda)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau\lambda} R_\xi(\tau) d\tau, & \lambda\in \boldsymbol{R}_1 \end{cases} \tag{1.81}$$

推论 在定理1.8条件下，如果过程还是实的，则

(a) $f_\xi(\lambda)=f_\xi(-\lambda), \quad \lambda\in R_1$;

(b) $$\begin{cases} R_\xi(\tau)=2\int_0^{\infty} f_\xi(\lambda)\cos\tau\lambda d\lambda, & \tau\in \boldsymbol{R}_1, \\ f_\xi(\lambda)=\frac{1}{\pi}\int_0^{\infty} R_\xi(\tau)\cos\tau\lambda d\tau, & \lambda\in \boldsymbol{R}_1. \end{cases} \tag{1.82}$$

4° 常见的相关函数和谱密度

以下我们举一些在理论上和科技应用中常遇到的相关函数和谱密度。

例1.7 设$\{\xi_n, n=0,\pm1,\cdots\}$为白噪声列(非标准)。$E\xi_k=0$，$E\xi_k\xi_l=\sigma^2\delta_{k,l}$，则

$$R_\xi(n)=\begin{cases}\sigma^2, & n=0; \\ 0, & n\neq 0.\end{cases}$$

从而$\sum_n |R_\xi(n)|<+\infty$，其谱密度$f_\xi(\lambda)$由(1.80)是

$$f_\xi(\lambda)=\frac{1}{2\pi}\left\{R_\xi(0)+2\sum_{k=1}^{\infty} R_\xi(k)\cos k\lambda\right\}=\frac{\sigma^2}{2\pi}, \quad -\pi\leqslant\lambda\leqslant\pi. \tag{1.83}$$

即谱密度是均匀分布的①。

例1.8 称$\{\xi_n, n=0,\pm1,\cdots\}$为平稳 Markov 序列，若存在$0<r<1$，使$\xi_n$的相关函数$R_\xi(n)$可表为

$$R_\xi(n)=\sigma^2 r^{|n|}, \quad 0<r<1, \quad n=0,\pm1,\cdots. \tag{1.84}$$

$$\sum_k |R_\xi(k)|\leqslant 2\sigma^2\sum_{k=0}^{\infty} r^k=2\sigma^2/(1-r)<+\infty,$$

---

① 信号处理中往往借用物理学中"白光"的概念，在工程技术上将独立同分布的序列也比喻成"白"的，称白噪声 (White Noise)，若噪声的谱密度不是均匀的则称为有色噪声 (Coloured Noise)。

故由 Wiener–Хинчин 定理

$$f_{\xi}(\lambda)=\frac{\sigma^2}{2\pi}\left\{\sum_{k=0}^{\infty}(re^{-i\lambda})^k+\sum_{k=-\infty}^{0}(r^{|k|}e^{-ik\lambda})-1\right\}$$

$$=\frac{\sigma^2}{2\pi}\left\{\frac{1}{1-r\,e^{-i\lambda}}+\frac{1}{1-r\,e^{i\lambda}}-1\right\}$$

$$=\frac{(1-r^2)\sigma^2}{2\pi}\frac{1}{|1-re^{-i\lambda}|^2},\quad 0\leqslant\lambda\leqslant\pi. \tag{1.85}$$

其形如图1.6所示。

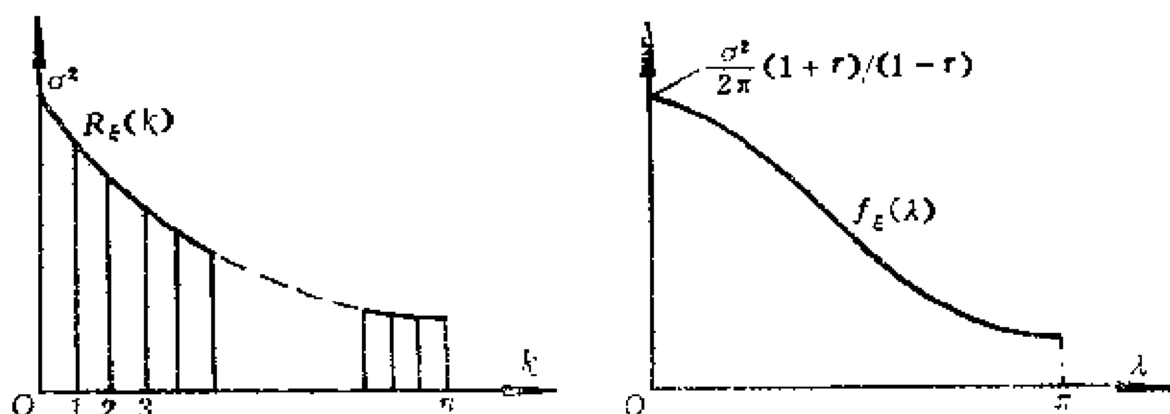


图 1.6

例1.9 白噪声过滤。设$\{\varepsilon_k\}$为白噪声列，令

$$\xi_n=\sum_{k=0}^{q}\theta_k\varepsilon_{n-k},\quad n=0,\pm1,\cdots,$$

则

$$R_{\xi}(n)=E\xi_{t+n}\xi_t=E\left(\sum_{k=0}^{q}\theta_k\varepsilon_{t+n-k}\right)\left(\sum_{l=0}^{q}\theta_l\varepsilon_{t-l}\right)$$

$$=\sum_{k=0}^{q}\sum_{l=0}^{q}\theta_k\theta_l\delta_{n-k,-l}. \tag{1.86}$$

不难看出$\sum_{n}|R_{\xi}(n)|<+\infty$，因而有谱密度

$$f_{\xi}(\lambda)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\left(\sum_{k,l=0}^{q}\theta_k\theta_l\delta_{n-k,-l}\right)e^{-in\lambda}$$

$$=\frac{1}{2\pi}\sum_{k,l=0}^{q}\theta_k\theta_l e^{-i(k-l)\lambda}$$

$$=\frac{1}{2\pi}\left|\sum_{k=0}^{q}\theta_k e^{-ik\lambda}\right|^2,\qquad -\pi\leqslant\lambda\leqslant\pi. \tag{1.87}$$

例1.10 随机电报信号．在例1.6中已得到随机电报信号ξ_t，$t\in\boldsymbol{R}_1$是平稳过程，它的协方差函数$\left(\text{均值 } E\xi_t=\frac{1}{2}\right)$

$$R_\xi(\tau)=\frac{1}{4}e^{-2a|\tau|},\quad \tau\in\boldsymbol{R}_1,\ a>0.$$

显见它是绝对可积的，因而谱密度$f_\xi(\lambda)$存在，可算得

$$f_\xi(\lambda)=\frac{1}{\pi}\int_0^\infty\left(\frac{1}{4}e^{-2a\tau}\right)\cos\tau\lambda d\tau$$

$$=\frac{1}{2\pi}\frac{a}{\lambda^2+4a^2},\qquad -\infty<\lambda<+\infty. \tag{1.88}$$

$R_\xi(\tau)$和$f_\xi(\lambda)$的图形和图1.6有类似之处。

其它常见的平稳过程的相关函数和谱密度可见于表1.1(见[9])。

表 1.1

相关函数 $R(\tau)$	谱密度 $f(\lambda)$
1. $\sigma^2 e^{-a\tau^2}$, $a>0$	1. $\frac{\sigma^2}{2\sqrt{\pi a}}e^{-\frac{\lambda^2}{4a}}$
2. $\sigma^2 e^{-a\tau^2}\cos\beta\tau$, $a>0$	2. $\frac{\sigma^2}{4\sqrt{\pi a}}\left[e^{-\frac{(\lambda+\beta)^2}{4a}}+e^{-\frac{(\lambda-\beta)^2}{4a}}\right]$
3. $\sigma^2 e^{-a|\tau|}\cos\beta\tau$, $a>0$	3. $\frac{\sigma^2 a}{2\pi}\left[\frac{1}{(\lambda-\beta)^2+a^2}+\frac{1}{(\lambda+\beta)^2+a^2}\right]$
4. $Ae^{-a|\tau|}(1+a|\tau|)$	4. $\frac{2Aa^3}{\pi(\lambda^2+a^2)^2}$
5. $R(\tau)=\begin{cases}\sigma^2(1-\|\tau\|), & \|\tau\|\leqslant 1\\ 0, & \|\tau\|>1\end{cases}$	5. $\frac{\sigma^2}{2\pi}\left(\frac{\sin\frac{\lambda}{2}}{\frac{\lambda}{2}}\right)^2$

§3 平稳随机过程的谱展式

1° 两个 Hilbert 空间的同构对应

在例1.4中，讨论过形如

$$\xi(t)=\sum_{k=1}^{N}\xi_k e^{i\lambda_k t}, \qquad t\in T \tag{1.89}$$

的过程，其中 $E\xi_k\equiv 0, E\xi_k\xi_l=\sigma^2\cdot\delta_{k,l}, \{\lambda_k\}$ 不相等，则 $\xi(t)$是平稳过程。以下我们将证明每一个平稳过程($E\xi(t)\equiv 0$) 可以看为类似(1.89)形式

$$\sum_k e^{i\lambda_k t}z(\Delta_k) \tag{1.90}$$

的(L^2)极限，其中$\{\lambda_k\}$不相同，$\{z(\Delta_k)\}$是随机变量，满足：

$$Ez(\Delta_k)\equiv 0,$$

$$Ez(\Delta_k)\overline{z(\Delta_j)}=0, \qquad \Delta_k\cap\Delta_j=\varnothing(k\neq j).$$

$\{\Delta_k\}$是一组互不相交的半开闭区间。这一结果称为平稳随机过程的谱展式。其确切意义和表达由下面给出，在此之前我们先证明一条引理，它对于随机过程的研究具有基本重要意义。

引理1.1 设$\{\xi_t, t=0,\pm1,\cdots\}$为平稳列，记

$$H_\xi=\mathscr{L}\{\xi_t, t=0,\pm1,\cdots\},$$

$$L^2(dF_\xi)=\left\{\varphi(\lambda):\ \varphi \text{ 是}[-\pi,\pi]\text{上的复值函数},\ \int_{-\pi}^{\pi}|\varphi|^2 dF_\xi<+\infty\right\}.\text{①}$$

则存在一个 $L^2(dF_\xi)$到 H_ξ 的等距对应映射 $\mathscr{K}$，在此对应下 H_ξ 与 $L^2(dF_\xi)$是两个同构的 Hilbert 空间。

证明 首先我们建立 $L^2(dF_\xi)\to H_\xi$ 的等距映射$\mathscr{K}$。

(1) 对 $e^{i\lambda t}\in L^2(dF)$，取$\mathscr{K}(e^{i\lambda t})=\xi_t$；

① 在$L^2(dF_\xi)$中，元素间的内积定义为$(\varphi,\psi)=\int_{-\pi}^{\pi}\varphi\bar{\psi}dF_\xi$。

(2) 对 $\varphi\in L\{e^{i\lambda t}\}$(线性集)，设 $\varphi=\sum_{1}^{n}c_k e^{i\lambda t_k}$，则

$$\mathscr{K}(\varphi)=\sum_{1}^{n}c_k\xi_{t_k}.$$

于是，对

$$\varphi_1=\sum_{1}^{n}\alpha_k e^{it_k\lambda},\quad \varphi_2=\sum_{1}^{m}\beta_l e^{is_l\lambda}\in L\{e^{i\lambda t}\},$$

有

$$\begin{aligned}(\mathscr{K}(\varphi_1),\mathscr{K}(\varphi_2))_{H_\xi}&=\left(\sum_{k=1}^{n}\alpha_k\xi_{t_k},\sum_{l=1}^{m}\beta_l\xi_{s_l}\right)_{H_\xi}\\&=\sum_{k=1}^{n}\sum_{l=1}^{m}\alpha_k\bar{\beta}_l(\xi_{t_k},\xi_{s_l})_{H_\xi}=\sum_{k=1}^{n}\sum_{l=1}^{m}R_\xi(t_k-s_l)\alpha_k\bar{\beta}_l\\&=\sum_{k=1}^{n}\sum_{l=1}^{m}\alpha_k\bar{\beta}_l\int_{-\pi}^{\pi}e^{i(t_k-s_l)\lambda}\,dF_\xi(\lambda)\\&=\int_{-\pi}^{\pi}\left(\sum_{k=1}^{n}\alpha_k e^{it_k\lambda}\right)\overline{\left(\sum_{l=1}^{m}\beta_l e^{is_l\lambda}\right)}dF_\xi(\lambda)=(\varphi_1,\varphi_2)_{L^2}.\end{aligned}\tag{1.91}$$

表明 $L\{e^{i\lambda t}\}$到 H_ξ 内的 $L\{\xi_t\}$是保内积的，因而也是"保距"的：

$$\|\varphi_1-\varphi_2\|_{L^2}^2=(\varphi_1-\varphi_2,\varphi_1-\varphi_2)=\|\mathscr{K}(\varphi_1)-\mathscr{K}(\varphi_2)\|_{H_\xi}^2.\tag{1.92}$$

因此，在 $L\{e^{i\lambda t}\}$中，$\mathscr{K}$将不同的元素对应于不同的 $L\{\xi_t\}$中的元素，由此得知$\mathscr{K}$在 $L\{\xi_t\}$有逆(可定义$\mathscr{K}^{-1}$)。此外，$\mathscr{K}$也还是线性的，即 $\varphi_j(j=1,2,\cdots,n)\in L\{e^{i\lambda t}\}$，则

$$\mathscr{K}\left(\sum_{j=1}^{n}\alpha_j\varphi_j\right)=\sum_{j=1}^{n}\alpha_j\mathscr{K}(\varphi_j).\tag{1.93}$$

(3) 考虑$L\{e^{i\lambda t}\}$元素列的极限点.设$\varphi_n\in L\{e^{i\lambda t}\}$是$L^2(dF_\xi)$中的一串 Cauchy 列，其极限点为 φ。由(1.92)等距对应关系知

$$\|\mathscr{K}(\varphi_m)-\mathscr{K}(\varphi_n)\|_{H_\xi}^2=\|\varphi_m-\varphi_n\|_{L^2}^2\to 0\quad(n,m\to\infty).$$

表明$\{\mathscr{K}(\varphi_n)\}$在 H_ξ 中也是 Cauchy 列，由完备性必存在 $\eta\in H_\xi$ 使

$$\eta=\lim_{n}\mathscr{K}(\varphi_n),\quad \varphi\in L^2(dF_\xi).\tag{1.94}$$

于是定义

$$\eta=\mathscr{K}(\varphi).\tag{1.95}$$

这样定义的对应关系仍然是保持内积的：设 $\varphi,\psi\in L^2(\mathrm{d}F_\xi)$，$\varphi_n$，$\psi_m\in L\{e^{i\lambda t}\}$使 $\varphi_n\xrightarrow{L^2}\varphi$，$\psi_m\xrightarrow{L^2}\psi$，则

$$
\begin{aligned}
(\mathscr{K}(\varphi),\ \mathscr{K}(\psi))_{H_\xi} &= (\lim_n \mathscr{K}(\varphi_n),\ \lim_m \mathscr{K}(\psi_m))_{H_\xi} \\
&= \lim_n \lim_m (\mathscr{K}(\varphi_n),\ \mathscr{K}(\psi_m))_{H_\xi} = \lim_n \lim_m (\varphi_n,\psi_m)_{L^2} \\
&= (\varphi,\psi)_{L^2}. \qquad (1.96)
\end{aligned}
$$

以上表明$\mathscr{K}$可以由 $L\{e^{i\lambda t}\}$扩充到$L^2(\mathrm{d}F_\xi)$并保持内积(且保距)，有逆映射$\mathscr{K}^{-1}$及线性性质等，因而 $L^2(\mathrm{d}F_\xi)$和 H_ξ 是两个同构的Hilbert空间。 ▎

定义1.8 设$\{Z(\lambda),\lambda\in\Lambda\}$是复值二阶矩随机过程($\Lambda$为区间)。

(1) $EZ(\lambda)\equiv 0$，　$\lambda\in\Lambda$； (1.97)

(2) 对于任意

$$\lambda_1<\lambda_2\leqslant\lambda_3<\lambda_4,\quad \lambda_i\in\Lambda,\ i=1,2,3,4, \qquad (1.98)$$

皆有

$$E(Z(\lambda_2)-Z(\lambda_1))\overline{(Z(\lambda_4)-Z(\lambda_3))}=0, \qquad (1.99)$$

则称$\{Z(\lambda)\}$为具有正交增量的随机过程。

引理1.2 设$\{\xi_t,t=0,\pm1,\cdots\}$是复值平稳序列，则存在一正交增量过程$\{Z(\lambda),\lambda\in[-\pi,\pi]\}$使得

$$E|Z(\lambda_2)-Z(\lambda_1)|^2=F_\xi(\lambda_2)-F_\xi(\lambda_1),\quad -\pi\leqslant\lambda_1<\lambda_2\leqslant\pi. \qquad (1.100)$$

其中 $F_\xi(\lambda)$为 ξ_t 的谱函数。

证明 利用引理1.1的结果知$\mathscr{K}$是 $L^2(\mathrm{d}F_\xi)$到 H_ξ 的等距映射，对$[a,b)\subset[-\pi,\pi]$，其示性函数

$$\chi_{[a,b)}(\lambda)\in L^2(\mathrm{d}F_\xi).$$

令

$$Z[a,b)=\mathscr{K}(\chi_{[a,b)}(\lambda))\in H_\xi, \qquad (1.101)$$

$$Z(\lambda)=\mathscr{K}(\chi_{[-\pi,\lambda)}),\qquad -\pi<\lambda\leqslant\pi, \qquad (1.102)$$

则由$\mathscr{K}$之线性性质

$$
\begin{aligned}
Z[a,b) &= \mathscr{K}(\chi_{[a,b)}) \\
&= \mathscr{K}(\chi_{[-\pi,b)} - \chi_{[-\pi,a)}) \\
&= \mathscr{K}(\chi_{[-\pi,b)}) - \mathscr{K}(\chi_{[-\pi,a)}) \\
&= Z(b) - Z(a). \qquad (1.103)
\end{aligned}
$$

验证正交增量的二条件：

(1) $E(Z(\lambda)) \equiv 0$, $-\pi < \lambda \leqslant \pi$.

这是显然的，因为$L\{\xi_t\}$中的元素皆为零均值[①]，因而H_ξ中的元素亦然。

(2) 对$-\pi < \lambda_1 < \lambda_2 \leqslant \lambda_3 < \lambda_4 \leqslant \pi$，由(1.103)

$$
\begin{aligned}
(Z[\lambda_1,\lambda_2),\ Z[\lambda_3,\lambda_4))_{H_\xi} &= (\mathscr{K}(\chi_{[\lambda_1,\lambda_2)}),\ \mathscr{K}(\chi_{[\lambda_3,\lambda_4)}))_{H_\xi} \\
&= (\chi_{[\lambda_1,\lambda_2)}, \chi_{[\lambda_3,\lambda_4)})_{L^2} \\
&= \int_{[\lambda_1,\lambda_2)\cap[\lambda_3,\lambda_4)} dF_\xi(\lambda) \\
&= 0. \qquad (1.104)
\end{aligned}
$$

此外 $\lim\limits_{\lambda\to-\pi+0} Z(\lambda) = \alpha$ 是存在的。

首先指出，对$-\pi < \lambda_1 < \lambda_2 \leqslant \pi$

$$
\begin{aligned}
\|Z(\lambda_2) - Z(\lambda_1)\|^2 &= \|\mathscr{K}(\chi_{[\lambda_1,\lambda_2)})\|^2_{H_\xi} \\
&= \|\chi_{[\lambda_1,\lambda_2)}\|^2_{L^2} \\
&= F_\xi(\lambda_2) - F_\xi(\lambda_1) \geqslant 0. \qquad (1.105)
\end{aligned}
$$

故

$$
\|Z(\lambda_2) - Z(\lambda_1)\|^2 \to 0, \qquad \lambda_1, \lambda_2 \to -\pi + 0.
$$

可见

$$
Z(\lambda) \to \alpha \in H_\xi, \qquad 当\ \lambda \to -\pi + 0.
$$

令$Z(-\pi) = \alpha$, $EZ(-\pi) = 0$, $\{Z(\lambda), -\pi \leqslant \lambda \leqslant \pi\}$是正交增量过程。由$F_\xi(\lambda)$性质知$Z(\lambda)$也是$(L_2)$左连续的。 ▎

以后我们都将$Z(\lambda)$标准化，认为$Z(-\pi) = 0$，否则由引理1.2的结果可令

---

① 前已交待我们通常都假定$E\xi(t) \equiv 0$。

$$\tilde{Z}(\lambda)=Z(\lambda)-\alpha, \qquad -\pi\leqslant\lambda\leqslant\pi,$$

则 $\tilde{Z}(\lambda)$ 是满足(1.100)的正交增量过程。

有了正交增量过程$\{Z(\lambda),-\pi\leqslant\lambda\leqslant\pi\}$，就可以按以下步骤定义随机积分

$$\int_{-\pi}^{\pi}\varphi(\lambda)\mathrm{d}Z(\lambda), \quad \varphi\in L^2(\mathrm{d}\,F_\xi).$$

(1) 设$[a,b)\subset[-\pi,\pi]$，$\chi_{[a,b)}\in L^2(\mathrm{d}\,F_\xi)$，定义

$$\int_{-\pi}^{\pi}\chi_{[a,b)}\,\mathrm{d}\,Z(\lambda)=Z[a,b)$$

(2) $\varphi=\sum_{k=1}^{n}a_k\chi_{I_k}\in L^2(\mathrm{d}\,F_\xi)$，其中$\{a_k\}$为常数，$\{I_k\}$是$n$个相互不相交的半开闭区间。则定义

$$\int_{-\pi}^{\pi}\varphi(\lambda)\mathrm{d}\,Z(\lambda)=\sum_{k=1}^{n}a_kZ(I_k).$$

不难看出，设$\varphi,\psi\in L\{\chi_{I_j}\}$，则有

$$\left(\int_{-\pi}^{\pi}\varphi\,\mathrm{d}\,Z(\lambda),\ \int_{-\pi}^{\pi}\psi\,\mathrm{d}\,Z(\lambda)\right)_{H_\xi}=(\varphi,\psi)_{L^2}.$$

事实上不妨认为

$$\begin{cases}\varphi(\lambda)=\alpha_j, & \text{当 } \lambda\in I_j=[a_j,b_j)\subset[-\pi,\pi];\\ \psi(\lambda)=\beta_j, & j=1,2,\cdots,n.\end{cases}$$

并且$\{I_j\}$是互不相交的，则

$$\left(\int_{-\pi}^{\pi}\varphi\mathrm{d}\,Z(\lambda),\int_{-\pi}^{\pi}\psi\mathrm{d}\,Z(\lambda)\right)_{H_\xi}=\left(\sum_j\alpha_jZ(I_j),\sum_l\beta_lZ(I_l)\right)_{H_\xi}$$

$$=\sum_{j,l}\alpha_j\bar{\beta}_l(Z(I_j),\ Z(I_l))_{H_\xi}=\sum_{j,l}\alpha_j\bar{\beta}_l(\chi_{I_j},\chi_{I_l})_{L^2}$$

$$=\left(\sum_j\alpha_j\chi_{I_j},\sum_l\beta_l\chi_{I_l}\right)_{L^2}$$

$$=\int_{-\pi}^{\pi}\varphi\bar{\psi}\,\mathrm{d}\,F_\xi=(\varphi,\psi)_{L^2}. \tag{1.106}$$

由此可知

$$\left\|\int_{-\pi}^{\pi}\varphi\mathrm{d}Z(\lambda)-\int_{-\pi}^{\pi}\psi\mathrm{d}Z(\lambda)\right\|_{H_\xi}^2=\|\varphi-\psi\|_{L^2}^2 \tag{1.107}$$

对 $\varphi,\psi\in L\{\chi_{I_j}\}$线性集成立。

(3) 设 $\varphi\in L^2(\mathrm{d}F_\xi)$，则存在$\varphi_n\in L\{\chi_{I_j}\}$， $\varphi_n\xrightarrow{L^2}\varphi$，并$\{\varphi_n\}$是 Cauchy 列。由(1.107)则

$$\left\|\int_{-\pi}^{\pi}\varphi_n\mathrm{d}Z(\lambda)-\int_{-\pi}^{\pi}\varphi_m\mathrm{d}Z(\lambda)\right\|_{H_\xi}^2=\|\varphi_n-\varphi_m\|_{L^2}^2\to 0\ (n,m\to\infty),$$

从而在 H_ξ 中必存在一个元素 ζ 使

$$\left\|\int_{-\pi}^{\pi}\varphi_n\mathrm{d}Z(\lambda)-\zeta\right\|_{H_\xi}^2\to 0\ \ (\text{当 } n\to\infty).$$

由此定义

$$\zeta=\int_{-\pi}^{\pi}\varphi\mathrm{d}Z(\lambda). \tag{1.108}$$

不难看出，这样的随机积分的定义在 H_ξ 中是一义的。对 $L^2(\mathrm{d}F_\xi)$中的不同元素 φ,ψ，由(1.108)所对应的 ζ_φ,ζ_ψ亦不相同，因为

$$(\zeta_\varphi,\zeta_\psi)_{H_\xi}=\left(\int_{-\pi}^{\pi}\varphi\mathrm{d}Z(\lambda),\int_{-\pi}^{\pi}\psi\mathrm{d}Z(\lambda)\right)_{H_\xi}=(\varphi,\psi)_{L^2},$$

$$\|\zeta_\varphi-\zeta_\psi\|_{H_\xi}^2=\|\varphi-\psi\|_{L^2}^2, \tag{1.109}$$

并且由(1.108)定义的随机积分还是线性的。

以下来证明平稳随机序列的谱展式。为简单起见，今后用如下记号①

$$\Pi\triangleq[-\pi,\pi],$$

$\mathscr{B}(\Pi)$表示$[-\pi,\pi]$上的 Borel 集系。

定理1.9 设$\{\xi_t,t=0,\pm1,\cdots\}$为平稳列，则 ξ_t 可表为

① 以后也用 $\boldsymbol{R}_1$表$(-\infty,\infty)$，$\mathscr{B}_1$表 $\boldsymbol{R}_1$ 上的 Borel 集系。

$$\xi_t=\int_{\Pi}\mathrm{e}^{\mathrm{i}\lambda t}\mathrm{d}Z_{\xi}(\lambda),\qquad t=0,\pm1,\cdots,\tag{1.110}$$

其中$\{Z_{\xi}(\lambda),\lambda\in\Pi\}$是标准化的具有正交增量的，左$(L^2)$连续的随机过程，并且

$$E(Z_{\xi}(\Delta_1)\cdot\overline{Z_{\xi}(\Delta_2)})=F_{\xi}(\Delta_1\cap\Delta_2),\ \Delta_1,\Delta_2\in\mathscr{B}(\Pi).\tag{1.111}$$

这样的正交增量过程唯一地由ξ_t所确定。

证明 由引理1.1和1.2知存在$\{Z(\lambda),\lambda\in\Pi\}$是标准化、左$(L^2)$连续(见(1.100))的正交增量过程，对$\varphi\in L^2(\mathrm{d}F_{\xi})$可定义随机积分

$$\zeta_{\varphi}=\int_{\Pi}\varphi\mathrm{d}Z(\lambda).$$

取$\varphi_n=\sum\limits_{j=1}^{n}C_j^{(n)}\chi_{I_j^{(n)}}\in L\{\chi_{I_j}\}$, $\varphi_n\xrightarrow{L^2}\varphi$，则

$$\begin{aligned}(\zeta_{\varphi},\xi_t)&=\left(\int_{\Pi}\varphi\mathrm{d}Z(\lambda),\xi_t\right)=\lim_n\left(\int_{\Pi}\varphi_n\mathrm{d}Z(\lambda),\xi_t\right)\\&=\lim_n\sum_{j=1}^{n}c_j^{(n)}\left(\int_{\Pi}\chi_{I_j^{(n)}}\mathrm{d}Z(\lambda),\xi_t\right)\\&=\lim_n\sum_{j=1}^{n}c_j^{(n)}(Z(I_j^{(n)}),\xi_t)\\&=\lim_n\sum_{j=1}^{n}c_j^{(n)}(\mathscr{H}(I_j^{(n)}),\mathscr{H}(\mathrm{e}^{\mathrm{i}\lambda t}))_{H_{\xi}}.\end{aligned}\tag{1.112}$$

(1.112)式是由(1.101)及引理1.1的$\mathscr{H}$定义得来的，再利用内积对应关系可得

$$\begin{aligned}(\zeta_{\varphi},\xi_t)&=\lim_n\sum_{j=1}^{n}c_j^{(n)}\int_{\Pi}\chi_{I_j^{(n)}}\mathrm{e}^{-\mathrm{i}\lambda t}\mathrm{d}F_{\xi}\\&=\int_{\Pi}\varphi\mathrm{e}^{-\mathrm{i}\lambda t}\mathrm{d}F_{\xi}.\end{aligned}\tag{1.113}$$

因此，考查

$$\begin{aligned}\left\|\xi_t-\int_{\Pi}\mathrm{e}^{\mathrm{i}\lambda t}\mathrm{d}Z(\lambda)\right\|^2&=\|\xi_t-\zeta_{\mathrm{e}^{\mathrm{i}\lambda t}}\|^2\\&=\|\xi_t\|^2+\|\zeta_{\mathrm{e}^{\mathrm{i}\lambda t}}\|^2-(\xi_t,\zeta_{\mathrm{e}^{\mathrm{i}\lambda t}})-(\zeta_{\mathrm{e}^{\mathrm{i}\lambda t}},\xi_t)\end{aligned}$$

$$= R_\xi(0) + \|e^{i\lambda t}\|_{L^2}^2 - \int_\Pi (e^{i\lambda t})(e^{-i\lambda t})\mathrm{d}F_\xi$$

$$- \int_\Pi (e^{-i\lambda t})(e^{i\lambda t})\mathrm{d}F_\xi$$

$$= R_\xi(0) + \int_\Pi \mathrm{d}F_\xi - \int_\Pi \mathrm{d}F_\xi - \int_\Pi \mathrm{d}F_\xi = 0,$$

对一切 $t = 0, \pm 1, \cdots$。表明

$$\xi_t = \int_\Pi e^{i\lambda t}\mathrm{d}Z(\lambda)。\tag{1.114}$$

以下证明$\{Z(\lambda)\}$之唯一性：

对$[a,b) \subset \Pi$，取一串 $\varphi_n = \sum_{j=1}^{n} c_j^{(n)} e^{it_j^{(n)}\lambda} \xrightarrow{L^2} \chi_{[a,b)}$，于是

$$Z[a,b) = Z(b) - Z(a) = \lim_n \int_\Pi \varphi_n(\lambda)\mathrm{d}Z(\lambda)$$

$$= \lim_n \sum_{j=1}^{n} c_j^{(n)} \xi_{t_j^{(n)}}。\tag{1.115}$$

表明 $Z[a,b) = Z(b) - Z(a)$是由$\{\xi_t\}$过程本身确定的，如果还有别的标准化正交增量$\{\zeta(\lambda), \lambda \in \Pi\}$也使得

$$\xi_t = \int_\Pi e^{i\lambda t}\mathrm{d}\zeta(\lambda),$$

则必然也可导致

$$\zeta[a,b) = \zeta(b) - \zeta(a) = \lim_n \sum_{j=1}^{n} c_j^{(n)} \xi_{t_j^{(n)}}$$

$$= Z(b) - Z(a)。\tag{1.116}$$

特别取$[-\pi, \lambda) = [a,b) \in \mathscr{B}(\Pi)$，则(1.116)变成

$$Z(\lambda) - Z(-\pi) = \zeta(\lambda) - \zeta(-\pi), \quad \lambda \in \Pi。$$

由标准化，则

$$Z(\lambda) = \zeta(\lambda), \qquad \lambda \in \Pi。 \quad |$$

我们对于平稳随机序列$\{\xi_t, t = 0, \pm 1, \cdots\}$谱展式中的标准正交增量过程$\{Z_\xi(\lambda), \lambda \in \Pi\}$的获得是通过等距映射$\mathscr{H}$来定义的，

即

(a) $Z_\xi^{(a)}(\lambda) \triangleq \mathscr{H}(\chi_{[-\pi,\lambda)})$。 (1.117)

然而许多数学工作者通过不同途径都可以等价地定义出 $Z_\xi(\lambda)$，以下这些定义都是富有启发性的。

(b) 设$\{\xi_t, t=0,\pm1,\cdots\}$为平稳列，对 $\lambda\in\Pi$，可以证明

$$\sum_{\substack{t=-N \\ t\neq 0}}^{N} \frac{e^{-it\lambda}}{it}\xi_t \tag{1.118}$$

当 $N\to\infty$ 在 H_ξ 中是有极限的，于是令

$$Z_\xi^{(b)}(\lambda) \triangleq \frac{1}{2\pi}\left\{\lambda\xi_0 - \sum_{\substack{t=-\infty \\ t\neq 0}}^{\infty} \frac{e^{-it\lambda}}{it}\xi_t\right\},\quad \lambda\in\Pi。 \tag{1.119}$$

可以证明它是正交增量过程，并且定义

$$\int_\Pi e^{i\lambda t}\,dZ_\xi^{(b)}(\lambda) \triangleq \lim_{\max|\Delta\lambda_k|\to 0} \sum_k e^{i\lambda_k t} Z_\xi^{(b)}(\Delta\lambda_k) \quad (H_\xi), \tag{1.120}$$

则

$$\xi_t = \int_\Pi e^{i\lambda t}\,dZ_\xi^{(b)}(\lambda)。$$

详细的证明可见[11]。

(c) 从统计的和工程科技应用分析观点出发，[12]给出如下 $Z_\xi(\lambda)$的产生方法：

设 $x_t, t=0,\pm1,\cdots,\pm N$ 是$\{\xi_t, t=0,\pm1,\cdots\}$的一段观测样本，作$\{x_t\}$的有限 Fourier 变换

$$d_\xi^{(N)}(\lambda) \triangleq \sum_{t=-N}^{N} x_t e^{-i\lambda t},\quad \lambda\in\Pi \tag{1.121}$$

并令

$$Z_N^{(c)}(\lambda) = \frac{1}{2\pi}\int_0^\lambda d_\xi^{(N)}(\mu)\,d\mu,\quad \lambda\in\Pi \tag{1.122}$$

则 $N\to\infty$时在 H_ξ 中有极限，记为 $Z_\xi^{(c)}(\lambda)$，则它是正交增量过

程并使 ξ_t 的谱展式为

$$\xi_t=\int_{\Pi}e^{i\lambda t}dZ_{\xi}^{(\varphi)}(\lambda).\qquad(1.123)$$

若 $\varphi\in L^2(dF_\xi)$，则定义随机积分

$$\int_{\Pi}\varphi(\lambda)dZ_{\xi}^{(\varphi)}(\lambda)=\lim_{M\to\infty}\frac{2\pi}{M}\sum_{n=0}^{M-1}\varphi(n\Delta_M)[Z_{\xi}^{(\varphi)}(-\pi+(n+1)\Delta_M)$$

$$-Z_{\xi}^{(\varphi)}(-\pi+n\Delta_M)]\ (\text{在 } H_\xi \text{ 中}).\qquad(1.124)$$

其中 M 为 $[-\pi,\pi]$ 的分割数，区间长为 $\Delta_M=\frac{2\pi}{M}$。不难看出(b)，(c)有共同之处。

(d) 从 Hilbert 空间酉算子的谱分解出发，[13]给出如下的观点：

在 H_ξ 中定义算子 U 为

$$U\xi_t=\xi_{t+1},\qquad(1.125)$$

则

$$(U\xi_t,U\xi_s)=(\xi_{t+1},\xi_{s+1})=R_\xi(t-s)$$

$$=(\xi_t,\xi_s).\qquad(1.126)$$

由此表明 U 是等距算子。可以证明 U 的定义域可扩张到 $H_\xi=\mathscr{L}\{\xi_t\}$，这种扩张不仅存在而且是唯一的并且还保持等距性。因此在 H_ξ 上引入了酉算子 U，根据 Hilbert 空间的算子理论[14]，必存在谱簇 $\{E_\lambda,\lambda\in\Pi\}$ 使得可将 U 表为

$$U=\int_{\Pi}e^{i\lambda}dE_\lambda,\qquad(1.127)$$

并且

$$U^n=\int_{\Pi}e^{in\lambda}dE_\lambda,\quad n=0,\pm1,\cdots.$$

$$\xi_n=U^n(\xi_0)=\int_{\Pi}e^{in\lambda}dE_\lambda\xi_0,\quad n=0,\pm1,\cdots.\qquad(1.128)$$

令

$$Z_\xi(\lambda)=E_\lambda\xi_0,\qquad \lambda\in\Pi. \tag{1.129}$$

则它是 L^2 左连续的正交增量过程，并有

$$F_\xi(\lambda)=\|E_\lambda\xi_0\|^2_{H_\xi}=(E_\lambda\xi_0,\xi_0)_{H_\xi},\qquad \lambda\in\Pi. \tag{1.130}$$

它是$\{\xi_n\}$的谱函数，左连续，非降。

对 $\varphi\in L^2(\mathrm{d}F_\xi)$，可自然给出随机积分定义

$$\eta_\varphi=\int_\Pi\varphi\mathrm{d}E_\lambda\xi_0,\qquad \varphi\in L^2(\mathrm{d}F_\xi) \tag{1.131}$$

并且这种 $L^2(\mathrm{d}F_\xi)$到 H_ξ 的映射是保内积(保距)的

$$\begin{aligned}(\eta_\varphi,\eta_\psi)&=\left(\int_\Pi\varphi\mathrm{d}E_\lambda\xi_0,\int_\Pi\psi\mathrm{d}E_\lambda\xi_0\right)_{H_\xi}\\&=\int_\Pi\varphi\overline{\psi}\mathrm{d}\|E_\lambda\xi_0\|^2=\int_\Pi\varphi\overline{\psi}\mathrm{d}F_\xi\\&=(\varphi,\psi)_{L^2}.\end{aligned} \tag{1.132}$$

(a),(b),(c),(d)的四种叙述方法对不同的读者可以各自选择，在自己知识范围内哪一种好理解就接受哪一种，并可以去参看相应的文献。

以上讨论的是离散足标(序列)的场合，在连续参数的过程下，也有类似的谱展式成立(见[1],[2])。

定理1.10 设$\{\xi_t,t\in \boldsymbol{R}_1\}$是均方连续的平稳过程，则必存在左均方连续的正交增量过程$\{Z(\lambda),\lambda\in\boldsymbol{R}_1\}$满足

(a) $EZ(\lambda)\equiv 0,\qquad \lambda\in\boldsymbol{R}_1$;

(b) 对 $\lambda_1<\lambda_2\leqslant\lambda_3<\lambda_4$,

$$E(Z(\lambda_2)-Z(\lambda_1))\overline{(Z(\lambda_4)-Z(\lambda_3))}=0;$$

(c) $\lim\limits_{\lambda\to-\infty}Z(\lambda)=a$ 存在 (H_ξ)

并表

$$\xi_t=\int_{\boldsymbol{R}_1}\mathrm{e}^{\mathrm{i}\lambda t}\mathrm{d}Z(\lambda),\qquad t\in\boldsymbol{R}_1. \tag{1.133}$$

2° 一般正交测度的随机积分

以上我们介绍了平稳过程$\{\xi_t,t\in T\}$(T或为 $\boldsymbol{R}_1$ 或为离散集)

可引出正交增量过程$\{Z_\xi(\lambda),\lambda\in\Lambda\}$，使得：

(1) $EZ_\xi(\Delta)\equiv 0$，$\Delta\in\mathscr{B}(\Lambda)$($\Lambda$上的Borel集系)；

(2) 对$\Delta_1,\Delta_2\in\mathscr{B}(\Lambda)$，则

$$EZ_\xi(\Delta_1)\overline{Z_\xi(\Delta_2)}=F_\xi(\Delta_1\cap\Delta_2).$$

其中$\Delta=[a,b)\subset\Lambda$时，$Z_\xi(\Delta)=Z_\xi(b)-Z_\xi(a)$。在引入上述随机集函数的基础上可逐步引入随机积分，事实上我们可以一般性地来讨论这一问题。

定义1.9 设$(\Lambda,\mathscr{B}(\Lambda),F)$是测度空间，$F(\Lambda)<+\infty$。如果在$\mathscr{B}(\Lambda)$上定义了取值在$H$的随机集函数[1]$Z(\cdot)$，$\cdot\in\mathscr{B}(\Lambda)$满足

$$(Z(A_1),Z(A_2))_H=F(A_1\cap A_2),\quad A_1,A_2\in\mathscr{B}(\Lambda),\tag{1.134}$$

则称Z是$(\Lambda,\mathscr{B}(\Lambda),F)$上的正交随机测度。

在平稳列的场合，$\Lambda=[-\pi,\pi]=\Pi$，$\mathscr{B}(\Lambda)=\mathscr{B}(\Pi)$，$F$即是$\xi_t$的谱测度。

在定义1.9中称$Z(\cdot)$，$\cdot\in\mathscr{B}(\Lambda)$为正交随机测度，是因为它具有$H$意义下的完全可加性。事实上可证明

定理1.11 设$Z(\cdot)$是$(\Lambda,\mathscr{B}(\Lambda),F)$上的随机集函数，则它是正交随机测度的充分必要条件是：

(a) $\|Z(A)\|_H^2=F(A)$，$A\in\mathscr{B}(\Lambda)$；　(1.135)

(b) 如$A_1\cap A_2=\varnothing$，$A_i\in\mathscr{B}(\Lambda)$，则

$$(Z(A_1),Z(A_2))_H=0;\tag{1.136}$$

(c) 若$A_i\in\mathscr{B}(\Lambda),i=1,2,\cdots,A_i\cap A_j=\varnothing\ (i\neq j)$，则在$H$中有

$$Z\Big(\bigcup_n A_n\Big)=\sum_n Z(A_n).\tag{1.137}$$

证明 先证必需性。由定义1.9的(1.134)式知(a),(b)两性质显然。完全可加性质(c)可由F的性质得出：令$A=\bigcup_n A_n$，

① 指对每个$s\in\mathscr{B}(\Lambda)$皆对应于具有二阶矩的随机变量$Z(s)$。

则

$$
\begin{aligned}
&\left\|Z(A)-\sum_{n=1}^{m}Z(A_n)\right\| \\
&= \|Z(A)\|^2+\sum_{n=1}^{m}\|Z(A_n)\|^2-\sum_{n=1}^{m}(Z(A),Z(A_n)) \\
&\quad -\sum_{n=1}^{m}(Z(A_n),Z(A)) \\
&= F(A)+\sum_{n=1}^{m}F(A_n)-2\sum_{n=1}^{m}F(A\cap A_n) \\
&= F(A)-\sum_{n=1}^{m}F(A_n)\to 0 \quad (m\to\infty).
\end{aligned}
$$

再证充分性。由性质(c)，对 $A_1,A_2\in\mathscr{B}(\Lambda)$有

$$
Z(A_1)=Z(A_1\cap A_2)+Z(A_1\backslash A_2),
$$
$$
Z(A_2)=Z(A_1\cap A_2)+Z(A_2\backslash A_1).
$$

再由性质(a),(b)可得

$$
\begin{aligned}
(Z(A_1),Z(A_2)) &= \|Z(A_1\cap A_2)\|^2+(Z(A_1\cap A_2),Z(A_2\backslash A_1)) \\
&\quad +(Z(A_1\backslash A_2),Z(A_1\cap A_2)) \\
&\quad +(Z(A_1\backslash A_2),\ Z(A_2\backslash A_1)) \\
&= F(A_1\cap A_2),
\end{aligned}
$$

对任意的 $A_1,A_2\in\mathscr{B}(\Lambda)$成立。由定义1.9知 $Z(\cdot)$是正交随机测度。 ∎

在引入了正交随机测度的基础上，我们可以引进随机积分如下（记 $\mathrm{d}F$ 所对应的 L^2 空间为 $L^2(\mathrm{d}F)$）：

(1) 对 $A\in\mathscr{B}(\Lambda)$，$\chi_A(\lambda)$，$\lambda\in\Lambda$ 为示性函数，则定义

$$
\int_\Lambda \chi_A(\lambda)\mathrm{d}Z(\lambda)=Z(A)\in H. \tag{1.138}
$$

(2) 对 $\varphi\in L\{\chi_{A_k}\}$，设为 $\varphi(\lambda)=\sum_{k=1}^{n}c_k\chi_{A_k}(\lambda)$，则定义

$$\int_\Lambda \varphi(\lambda)\mathrm{d}Z(\lambda) = \sum_{k=1}^{n} c_k Z(A_k) \in H, \tag{1.139}$$

并且对 $\varphi, \psi \in L\{\chi_{A_i}\}$，有

$$\left(\int_\Lambda \varphi \mathrm{d}Z, \int_\Lambda \psi \mathrm{d}Z\right)_H = \sum_{k,l} c_k \bar{d}_k (Z(A_k), Z(B_l))_H$$

$$= \left(\sum_k c_k \chi_{A_k}, \sum_l d_l \chi_{B_l}\right)_{L^2}$$

$$= (\varphi, \psi)_{L^2} = \int_\Lambda \varphi(\lambda)\overline{\psi(\lambda)}\mathrm{d}F(\lambda) \tag{1.140}$$

及

$$\left\|\int_\Lambda \varphi \mathrm{d}Z - \int_\Lambda \psi \mathrm{d}Z\right\|_H^2 = \|\varphi - \psi\|_{L^2}^2. \tag{1.141}$$

若记

$$L(Z) = \left\{\int_\Lambda \varphi \mathrm{d}Z;\ \varphi \in L\{\chi_{A_i}\}\right\} \subset H, \tag{1.142}$$

则(1.140)，(1.141)表明 $L(Z)$和 $L\{\chi_{A_i}\}$之间有等距对应关系。

(3) 对 $\varphi \in L^2(\mathrm{d}F)$，取一串 $\varphi_n \in L\{\chi_{A_i}\}$，$\varphi_n \xrightarrow{L^2} \varphi$，则$\{\varphi_n\}$所对应的随机积分

$$I(\varphi_n) = \int_\Lambda \varphi_n(\lambda)\mathrm{d}Z(\lambda), \qquad n = 1, 2, \cdots$$

也是H中的 Cauchy 列，这可由(1.141)看出

$$\|I(\varphi_n) - I(\varphi_m)\|_H^2 = \|\varphi_n - \varphi_m\|_{L^2}^2 \to 0 \quad (n, m \to \infty).$$

由H之完备性，$\exists \zeta_\varphi \in H$ 使

$$\zeta_\varphi = \lim_n \int_\Lambda \varphi_n(\lambda)\mathrm{d}Z(\lambda).$$

和以往类似，定义

$$\zeta_\varphi = I(\varphi) = \int_\Lambda \varphi(\lambda)\mathrm{d}Z(\lambda),$$

不难证明上述随机积分是线性的，并建立了$(L(Z))^* = \mathscr{L}\{Z\}$和$L^2(\mathrm{d}F)$二个 Hilbert 空间彼此之间的等距对应关系。

需要指出的是：我们以上假定了$(\Lambda, \mathscr{B}(\Lambda), F)$测度空间是$F(\Lambda) < +\infty$，其实可放宽。对

$$\mathscr{B}^0(\Lambda) = \{A;\ A \in \mathscr{B}(\Lambda), F(A) < +\infty\}$$

定义随机集函数$Z(\cdot)$，$\cdot \in \mathscr{B}^0(\Lambda)$，若

$$(Z(A_1), Z(A_2)) = F(A_1 \cap A_2),\quad A_1, A_2 \in \mathscr{B}^0(\Lambda), \qquad (1.143)$$

就称为在$(\Lambda, \mathscr{B}(\Lambda), F)$上定义了正交随机测度$Z(\cdot)$，$\cdot \in \mathscr{B}(\Lambda)$。用同样方法可得到以上随机积分的定义及$\mathscr{L}\{Z\}$和$L^2(\mathrm{d}F)$的等距同构对应(见[2])。

3° Karhunen 定理及其应用

定理 1.9 及 1.10 讲的是平稳随机过程的谱展式，这些结果可推到一般的二阶矩过程，这对非平稳过程的研究具有重要意义。

定理1.12 设$\{\xi_t, t \in T\}$是具有二阶矩的过程$(E\xi_t \equiv 0)$，其协方差函数$E\xi_t\bar{\xi}_s = R(t,s)$可表为

$$R(t,s) = \int_\Lambda f(t,\lambda)\overline{f(s,\lambda)}\,\mathrm{d}F(\lambda),\qquad t,s \in T, \quad (1.144)$$

其中$\Lambda \in \mathscr{B}_1$，$\mathrm{d}F$是可测空间$(\Lambda, \mathscr{B}(\Lambda))$的有限测度，$F(\Lambda) < +\infty$，则必存在一个定义在$(\Lambda, \mathscr{B}(\Lambda), F)$上的随机正交测度$Z(\cdot)$，$\cdot \in \mathscr{B}(\Lambda)$使得

$$\xi_t = \int_\Lambda f(t,\lambda)\mathrm{d}Z(\lambda),\qquad t \in T. \qquad (1.145)$$

证明 本定理证明的手法都比较典型，和定理1.9及2°有相似之处，读者可按以下线索证明本定理。

1. 假定$L\{f(t,\lambda), t \in T\}$线性集在$L^2(\mathrm{d}F)$中是稠密系。于是：

(a) 构造随机正交测度$Z(\cdot)$，$\cdot \in \mathscr{B}(\Lambda)$。对$A \in \mathscr{B}(\Lambda)$，

$\chi_A(\lambda)\in L^2(\mathrm{d}F)$，则存在

$$\varphi_A^{(n)}(\lambda)=\sum_{k=1}^{n}c_k^{(n)}f(t_k^{(n)},\lambda)\in L\{f(t,\lambda),t\in T\}$$

L^2 收敛到 $\chi_A(\lambda)$。令

$$Z_n^{(A)}=\sum_{k=1}^{n}c_k^{(n)}\xi_{t_k^{(n)}},\qquad n=1,2,\cdots。$$

则

$$\|Z_n^{(A)}-Z_m^{(A)}\|_H^2=\|\varphi_A^{(n)}(\lambda)-\varphi_A^{(m)}(\lambda)\|_{L^2}^2\to 0\quad(n,m\to\infty)。$$

可见有 $Z(A)\in H$ 使 $Z_n^{(A)}\xrightarrow{H}Z(A)$，即

$$Z(A)\triangleq\lim_n Z_n^{(A)},\qquad A\in\mathscr{B}(\Lambda)。\qquad(1.146)$$

可以看出它是正交测度(并且 $EZ(A)=0$)。

(b) 对 $\varphi\in L^2(\mathrm{d}F)$ 可定义随机积分(见 2°)，于是对 $f(t,\lambda)\in L^2(\mathrm{d}F)$，令

$$\zeta_t=\int_\Lambda f(t,\lambda)\mathrm{d}Z(\lambda),\qquad t\in T,\qquad(1.147)$$

则

$$\|\zeta_t\|_H^2=\int_\Lambda|f(t,\lambda)|^2\mathrm{d}F=\|\xi_t\|_H^2。\qquad(1.148)$$

(c) 利用以下等式：对 $A\in\mathscr{B}(\Lambda)$，有

$$(\xi_t,Z(A))_H=\int_A f(t,\lambda)\mathrm{d}F(\lambda)。\qquad(1.149)$$

因为存在 $\varphi_n(\lambda)\in L\{\chi_{A_t}\}$ 使 $\varphi_n(\lambda)\xrightarrow{L^2}f(t,\lambda)$，从而由(1.149)可得

$$(\xi_t,\zeta_t)=\lim_n\left(\xi_t,\int_\Lambda\varphi_n(\lambda)\mathrm{d}Z(\lambda)\right)$$

$$=\lim_n\int_\Lambda f(t,\lambda)\overline{\varphi_n(\lambda)}\mathrm{d}F=\int_\Lambda|f(t,\lambda)|^2\mathrm{d}F。$$

可见

$$\|\xi_t-\zeta_t\|^2=\|\xi_t\|^2+\|\zeta_t\|^2-(\xi_t,\zeta_t)-(\zeta_t,\xi_t)=0,$$

此即 (1.145) 式。

2. 假定$L\{f(t,\lambda),t\in T\}$在 $L^2(\mathrm{d}F)$中不稠密。于是:

a. 补充$L\{f(t,\lambda),t\in T\}$使之是 $L^2(\mathrm{d}F)$的稠密集。为此只需考虑与一切$f(t,\lambda),t\in T$正交的全体$L^2(\mathrm{d}F)$中的元素$g(t',\lambda)$, $t'\in T',T\cap T'=\varnothing$, 并考虑如下的线性集

$$L\{f(t,\lambda),g(t',\lambda),\ t\in T,t'\in T'\},$$

则它在 $L^2(\mathrm{d}F)$中稠密。

b. 利用复 Gauss 存在定理(定理 1.3), 可取正态过程$\{\tilde{\xi}_t,t\in T'\}$使得

(i) $E\tilde{\xi}_t\equiv 0,\quad t\in T'$;

(ii) $R_{\tilde{\xi}}(t,s)=\int_\Lambda g(t,\lambda)\overline{g(s,\lambda)}\mathrm{d}F,\quad s,t\in T'$;

(iii) $\tilde{\xi}_t$与ξ_s是独立的$(t\in T',s\in T)$。

于是令

$$\zeta_t=\begin{cases}\xi_t, & \text{如 } t\in T;\\ \tilde{\xi}_t, & \text{如 } t\in T'.\end{cases}$$

则$\{\zeta_t,t\in T\cup T'\}$的协方差函数为

$$R_\zeta(t,s)=\int_\Lambda h(t,\lambda)\overline{h(s,\lambda)}\mathrm{d}F(\lambda),\quad t,s\in T\cup T',\quad(1.150)$$

其中

$$h(t,\lambda)=\begin{cases}f(t,\lambda), & t\in T;\\ g(t,\lambda), & t\in T'.\end{cases}\qquad(1.151)$$

在以上补充修正之后即可用已经证明的第一种情况的结论，故存在$(\Lambda,\mathscr{B}(\Lambda),F)$上的随机正交测度 $Z(\cdot),\cdot\in\mathscr{B}(\Lambda)$使得

$$\xi_t=\int_\Lambda h(t,\lambda)\mathrm{d}Z(\lambda),\quad t\in T\cup T'$$

成立。特别当 $t\in T$ 时即得本定理所需结论。∎

显然，定理1.9,1.10都可看为Karhunen定理的特例：

设 $\Lambda=[-\pi,\pi]$，$f(t,\lambda)=e^{it\lambda}$，$t=0,\pm1,\cdots,\xi_t$ 是平稳序列，相关函数有以下表达式

$$R(t,s)=R(t-s)=\int_\Lambda e^{it\lambda}\overline{e^{is\lambda}}\mathrm{d}F(\lambda)\quad,$$

则存在 $(\Lambda,\mathscr{B}(\Lambda),F)$ 上的随机正交测度 $Z(\cdot),\cdot\in\mathscr{B}(\Lambda)$ 使

$$\xi_t=\int_\Lambda e^{i\lambda t}\mathrm{d}Z(\lambda),\quad t=0,\pm1,\cdots.$$

若 $\Lambda=R_1,T=R_1,\xi_t$ 是平稳过程，则定理1.10亦可由Karhunen定理导出。

作为Karhunen定理的应用我们来介绍几个有用的结果

定理1.13 设 $\xi_t,a\leqslant t\leqslant b$ 是实均方连续的随机过程($E\xi_t\equiv 0$)，则可将 ξ_t 展成

$$\xi_t=\sum_{k=1}^{\infty}\frac{\varphi_k(t)}{\sqrt{\lambda_k}}z_k\ (\text{在}H\text{内}),\qquad(1.152)$$

其中 $\{z_k\}$ 为实随机变量，$Ez_k=0$，$Ez_kz_l=\delta_{k,l}$，λ_k 为积分方程

$$\varphi(t)=\lambda\int_a^b R(t,s)\varphi(s)\mathrm{d}s\qquad(1.153)$$

的特征根，$\varphi_k(t)$ 为相应的特征向量，$R(t,s)=E\xi_t\xi_s$。

证明 利用积分方程中的Mercer定理（参见[15]）可得

$$R(t,s)=\sum_{k=1}^{\infty}\frac{\varphi_k(t)\varphi_k(s)}{\lambda_k}\ (\lambda_k>0),\qquad(1.154)$$

其中 $\varphi_k(t)$，λ_k 由(1.153)式所决定。

令

$$f(t,\lambda)=\begin{cases}\dfrac{\varphi_k(t)}{\sqrt{\lambda_k}}k, & \text{当 } \lambda=k \ (k=1,2,\cdots);\\ 0, & \text{其它,}\end{cases} \tag{1.155}$$

并在$(R_1,\mathscr{B}_1)$上考虑测度F

$$F=\begin{cases}F\{k\}=\dfrac{1}{k^2}, & k=1,2,\cdots;\\ F\{A\}=0, & A\in\mathscr{B}_1, A\text{中不含正整数点.}\end{cases} \tag{1.156}$$

显然，$F\{R_1\}=\sum\limits_{k=1}^{\infty}\frac{1}{k^2}<+\infty$是有穷测度，并且由

$$\int_{R_1}|f(t,\lambda)|^2\mathrm{d}F=\sum_{k=1}^{\infty}\frac{|\varphi_k(t)|^2}{\lambda_k}<+\infty$$

知$f(t,\lambda)\in L^2(\mathrm{d}F)$，此外$R(s,t)$可表为

$$R(s,t)=\int_{R_1}(f(t,\lambda))(f(s,\lambda))\mathrm{d}F$$

$$=\sum_{k=1}^{\infty}\frac{\varphi_k(s)\varphi_k(t)}{\lambda_k}. \tag{1.157}$$

由 Karhunen 定理知，存在正交随机测度$Z(\cdot)$定义于$(R_1,\mathscr{B}_1,F)$使得ξ_t可表为

$$\xi_t=\int_{R_1}f(t,\lambda)\mathrm{d}Z(\lambda),\quad a\leqslant t\leqslant b. \tag{1.158}$$

由于

$$\|Z(A)\|^2=0 \quad (\text{当}A\text{不包含任何正整数点}),$$

故$Z(\cdot)$只在正整数点上有测度，因此(1.158)可写为

$$\xi_t=\sum_{k=1}^{\infty}\frac{\varphi_k(t)}{\sqrt{\lambda_k}}kZ(k),\quad a\leqslant t\leqslant b. \tag{1.159}$$

令

$$z_k=kZ(k),\quad k=1,2,\cdots.$$

则

$$E[z_k]=0\ (因为\ E\xi_t=0),$$

$$(z_k,z_l)_H=(kZ(k),lZ(l))_H=klF(\{k\}\cap\{l\})$$
$$=0,\qquad 当\ k\neq l.$$

可见$\{z_k\}$是不相关列。 ▌

定理1.14 设$\{\xi_t,t=0,\pm1,\cdots\}$是具有谱密度的平稳列，则$\xi_t$ 可展成(在H内)

$$\xi_t=\sum_{k=-\infty}^{\infty}a_k\varepsilon_{t-k},\qquad t=0,\pm1,\cdots,\qquad(1.160)$$

其中 $$E\varepsilon_t=0,\qquad E\varepsilon_t\varepsilon_s=\delta_{t,s}.$$

证明 相关函数

$$E\xi_t\bar{\xi}_s=R(t-s)=\int_\Pi e^{i(t-s)\lambda}f_\xi(\lambda)d\lambda$$
$$=\int_\Pi(e^{it\lambda}\sqrt{f_\xi(\lambda)})\overline{(e^{is\lambda}\sqrt{f_\xi(\lambda)})}d\lambda.$$

由 Karhunen 定理知，必存在正交随机测度$Z(\cdot)$，使得可将ξ_t表为

$$\xi_t=\int_\Pi e^{i\lambda t}\sqrt{f_\xi(\lambda)}\,dZ(\lambda).\qquad(1.161)$$

由于$\sqrt{f_\xi(\lambda)}\in L^2(d\lambda)$，故

$$\sqrt{f(\lambda)}=\sum_{k=-\infty}^{\infty}a_k\frac{e^{-ik\lambda}}{\sqrt{2\pi}}\ (L^2),\qquad(1.162)$$

即

$$\xi_t=\sum_{k=-\infty}^{\infty}a_k\left(\int_\Pi e^{i\lambda t}\frac{e^{-ik\lambda}}{\sqrt{2\pi}}dZ\right)$$
$$=\sum_{k=-\infty}^{\infty}a_k\varepsilon_{t-k},$$

其中

$$\varepsilon_t = \frac{1}{\sqrt{2\pi}}\int_{\Pi} e^{it\lambda} dZ(\lambda), \qquad t = 0, \pm 1, \cdots. \tag{1.163}$$

满足

$$E\varepsilon_t = 0,$$

$$(\varepsilon_t, \varepsilon_s) = \left(\frac{1}{\sqrt{2\pi}}\int_{\Pi} e^{it\lambda} dZ(\lambda), \frac{1}{\sqrt{2\pi}}\int_{\Pi} e^{is\lambda} dZ(\lambda)\right)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^2 \int_{\Pi} e^{i(t-s)\lambda}\, d\lambda = \delta_{t,s}. \quad \blacksquare$$

4° 随机过程的离散化

在前几节中我们介绍的结果，都以序列为重点，因为许多结果都相类似地可出现在连续足标的过程场合，今后在时间序列理论中都只讨论离散足标的序列.这样做，不仅在理论上可以更加简明，而更重要的是：随着电算机的飞速发展和广泛的普及，数据的分析和处理几乎都是离散化的，因为近代电算机的运算都是离散形式，那怕原来的对象是连续量也要经过 A/D 转换加以离散化，称为采样。但是采样必须是合理的才能保存原来的信息，否则可能在采样中丢失许多信息以致于面目全非。以下的采样定理虽然简单却是工程技术数值化方法的理论基础，被广泛地应用于各个领域。

定理1.15（采样定理） 设$\{\xi_t, t \in \boldsymbol{R}_1\}$是平稳过程，并其谱函数 $F(\lambda)$满足

$$\int_{|\lambda| \geq 2\pi w} dF(\lambda) = 0, \tag{1.164}$$

则

$$\xi_t = \sum_{n=-\infty}^{\infty} \xi_{(n/2w)} \frac{\sin(2\pi w t - n\pi)}{2\pi w t - n\pi}, \qquad t \in \boldsymbol{R}_1. \tag{1.165}$$

证明 由平稳过程的谱展式，存在定义在$(\boldsymbol{R}_1, \mathscr{B}_1, F)$上的正交随机测度 $Z(\cdot), \cdot \in \mathscr{B}_1$ 使得

$$\xi_t = \int_{R_1} e^{i\lambda t} dZ(\lambda), \quad t \in R_1. \tag{1.166}$$

然而利用(1.164)，当 $A \in \mathscr{B}_1$ 并且

$$A \subset \{(-\infty, -2\pi w) \cup (2\pi w, +\infty)\}$$

时

$$\left\| \int_A e^{i\lambda t} dZ(\lambda) \right\|_H^2 = \left\| \int_{R_1} \chi_A(\lambda) e^{i\lambda t} dZ(\lambda) \right\|_H^2$$

$$= \| \chi_A(\lambda) e^{i\lambda t} \|_{L^2}^2 = \int_{R_1} | \chi_A(\lambda) e^{i\lambda t} |^2 dF$$

$$= F(A) = 0. \tag{1.167}$$

故(1.166)实质上是

$$\xi_t = \int_{-2\pi w}^{2\pi w} e^{i\lambda t} dZ(\lambda), \quad t \in R_1. \tag{1.168}$$

对固定的 t，在 $[-2\pi w, 2\pi w]$ 上 $e^{i\lambda t}$ 可展成

$$e^{i\lambda t} \sim \sum_{n=-\infty}^{\infty} C_n(t) e^{in\frac{\lambda}{2w}}. \tag{1.169}$$

而

$$C_n(t) = \frac{1}{4\pi w} \int_{-2\pi w}^{2\pi w} e^{i\lambda t} e^{-in\frac{\lambda}{2w}} d\lambda$$

$$= \frac{1}{4\pi w} \int_{-2\pi w}^{2\pi w} e^{i\lambda\left(t - \frac{n}{2w}\right)} d\lambda$$

$$= \frac{\sin 2\pi w\left(t - \frac{n}{2w}\right)}{2\pi w\left(t - \frac{n}{2w}\right)} = \frac{\sin(2\pi wt - n\pi)}{2\pi wt - n\pi}.$$

由于 $e^{i\lambda t}$ 是 λ 的连续函数，且在 $[-2\pi w, 2\pi w]$ 上是有界变差的，所以其 Fourier 级数之部分和 $S_n(\lambda)$ 对 n 是一致有界的，且在 $(-2\pi w, 2\pi w)$ 上收敛于 $e^{i\lambda t}$（参见[8]）。由控制收敛定理

$$\lim_{N \to \infty} \int_{|\lambda| < 2\pi w} \left| e^{i\lambda t} - \sum_{n=-N}^{N} C_n(t) e^{in\frac{\lambda}{2w}} \right|^2 dF(\lambda) = 0.$$

可见

$$\left\|\xi_t - \sum_{n=-N}^{N} C_n(t)\xi_{(n/2w)}\right\|_H^2 \to 0 \quad (N\to\infty),$$

即在H中

$$\xi_t = \sum_{n=-\infty}^{\infty} \xi_{(n/2w)} \frac{\sin(2\pi wt - n\pi)}{2\pi wt - n\pi}. \tag{1.170}$$

定理至此证毕. ▌

此定理告诉我们:假若平稳过程$\xi_t, t\in \boldsymbol{R}_1$的谱$F(\lambda)$在$(-\infty, -2\pi w)$和$(2\pi w, +\infty)$取常数(分两段), 则$\xi_t$完全被可列个随机变量$\{\xi_{(n/2w)}\}$所确定, 也就是说$\{\xi_t\}$被间隔为$\Delta=\frac{1}{2w}$的采样序列$\eta_n=\xi_{(n/2w)}$所确定。应用中, 多假定$\xi_t$为有"有限带宽"的功率谱密度

$$f_\xi(\lambda)\equiv 0, \qquad |\lambda|>2\pi w. \tag{1.171}$$

而应用上述采样定理, 此时称w为过程ξ_t的"上界频"。w的确定在实用中可用多种途径解决, 或者由物理(工程)背景分析知道过程ξ_t的谐波成分不会超过w(Hz)或者由接收信号的仪器本身有一个上界截频w来确定。值得注意的是定理1.15只要求

$$\int_{|\lambda|>2\pi w} \mathrm{d}F(\lambda)=0$$

或通用条件(1.171),并没有说

$$f_\xi(\lambda)\neq 0, \quad |\lambda|\leqslant 2\pi w. \tag{1.172}$$

因此w频限只是个"上界频率", 因为许多场合并不能确切估计出信号ξ_t的频限临界值。因此估宽一点, 仍有(1.170)的展式成立。当然, 由于w估大了, 采样点$\{\xi_{(n/2w)}\}$就会增多(相同时间间隔内),工作量就会增加。但是,无论如何不允许采样间隔$\Delta<\frac{1}{2w}$(w为ξ_t的频限临界值)。

§4 平稳随机过程的强大数律

1° 关于数学期望的强大数律

时间序列的研究往往从一列过程(序列)的样本出发来估计一、二阶矩及谱函数或谱密度。作为经典的方法，这一节将介绍平稳过程的强大数定律，主要解决如何从一段观测样本出发来估计期望与相关函数。由于样本只有一次实现，并且不是相互独立的简单样本，读者将看到其证明方法与经典统计不一样。

定理1.16 设x_t是零均值平稳列，$R_x(k)$为其相关函数，若对充分大的n，有

$$\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)R_x(j)e^{-ij\lambda}=O\left(\frac{1}{n^{\alpha}}\right), \quad (1.173)$$

其中(1.173)对$\lambda\in\Pi$一致成立，$\alpha>0$，则① 对$\lambda\in\Pi$，

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_je^{-i\lambda j}=0, \quad \text{a.s.}. \quad (1.174)$$

证明 首先指出有以下等式成立，对$\lambda\in\Pi$，

$$\begin{aligned}E\left\{\left|\frac{1}{n}\sum_{j=1}^{n}x_je^{-ij\lambda}\right|^2\right\}&=\frac{1}{n^2}\sum_{k,j=1}^{n}R_x(k-j)e^{-i(k-j)\lambda}\\&=\frac{1}{n^2}\left(\sum_{l=1-n}^{0}\sum_{k=1}^{l+n}+\sum_{l=1}^{n-1}\sum_{k=l+1}^{n}\right)R_x(l)e^{-il\lambda}\\&=\frac{1}{n}\sum_{l=-(n-1)}^{n-1}\left(1-\frac{|l|}{n}\right)R_x(l)e^{-il\lambda}.\end{aligned}$$

(1.175)

(1) 首先证明$\lambda=0$时，(1.174)成立。由于(1.173)对某个

① 即概率为1地当n充分大有$\sum_{j=1}^{n}x_je^{-ij\lambda}=o(n)$，$\lambda\in\Pi$。

$\alpha>0$ 成立，故存在 $\beta>0$ 使 $\alpha\beta>1$。下面分以下两步证明：

A．对于 $n_m=[m^\beta]+1, m=1,2,\cdots$ 的正整数列$\{n_m\}$，由(1.175)，(1.173)可得

$$E\left\{\left|\frac{1}{n_m}\sum_{j=1}^{n_m}x_j\right|^2\right\}=\frac{1}{n_m}\sum_{l=-(n_m-1)}^{(n_m-1)}\left(1-\frac{|l|}{n_m}\right)R_x(l)$$

$$\leqslant\frac{K}{n_m^\alpha}\leqslant\frac{K}{(m^\beta)^\alpha}\quad(\alpha\beta>1).$$

对 $\varepsilon>0$，利用车贝谢夫不等式

$$\sum_{m=1}^{\infty}P\left\{\left|\frac{1}{n_m}\sum_{j=1}^{n_m}x_j(\omega)\right|\geqslant\varepsilon\right\}\leqslant\sum_{m=1}^{\infty}\frac{K}{\varepsilon^2m^{\alpha\beta}}<+\infty.$$

由 Borel-Cantelli 引理(参见[9],[16])，当 m 充分大时有

$$\left|\frac{1}{n_m}\sum_{j=1}^{n_m}x_j(\omega)\right|<\varepsilon,\quad \text{a.s.}.$$

由 ε 的任意性，表明

$$\lim_{m}\frac{1}{n_m}\sum_{j=1}^{n_m}x_j(\omega)=0,\quad \text{a.s.}. \tag{1.176}$$

B．对于一般 n 正整数，设 $n_m\leqslant n<n_{m+1}$，则有

$$E\left\{\max_{n_m\leqslant n<n_{m+1}}\left|\frac{1}{n}\sum_{j=1}^{n}x_j-\frac{1}{n}\sum_{j=1}^{n_m}x_j\right|^2\right\}$$

$$\leqslant\frac{1}{n_m^2}E\left\{\max_{n_m\leqslant n<n_{m+1}}\left|\sum_{j=n_m+1}^{n}x_j\right|^2\right\}$$

$$\leqslant\frac{1}{n_m^2}E\left\{\left(\sum_{j=n_m+1}^{n_{m+1}}|x_j|\right)^2\right\}$$

$$=\frac{1}{n_m^2}\sum_{k=n_m+1}^{n_{m+1}}\sum_{j=n_m+1}^{n_{m+1}}E(|x_j||x_k|)$$

$$\leqslant\frac{R_x(0)}{n_m^2}[n_{m+1}-(n_m+1)+1]^2. \tag{1.177}$$

然而

$$n_{m+1}-n_m=[(m+1)^\beta]-[m^\beta]$$
$$\leqslant(m+1)^\beta-(m^\beta-1)$$
$$=K_1m^{\beta-1}+o(m^{\beta-1})\quad(\because\beta>1),\ K_1\text{ 为常数},$$

故

$$\frac{(n_{m+1}-n_m)^2}{n_m^2}\leqslant K_2\frac{m^{2\beta-2}}{([m^\beta]+1)^2}\leqslant K_2\frac{m^{2\beta-2}}{m^{2\beta}}=\frac{K_2}{m^2}. \quad (1.178)$$

将(1.178)代入(1.177)得

$$E\left\{\max_{n_m\leqslant n<n_{m+1}}\left|\frac{1}{n}\sum_{j=1}^{n}x_j-\frac{1}{n}\sum_{j=1}^{n_m}x_j\right|^2\right\}\leqslant\frac{K_2}{m^2}.$$

因此

$$\sum_{m=1}^{\infty}P\left\{\max_{n_m\leqslant n<n_{m+1}}\left|\frac{1}{n}\sum_{1}^{n}x_j-\frac{1}{n}\sum_{1}^{n_m}x_j\right|\geqslant\varepsilon\right\}$$
$$\leqslant\frac{1}{\varepsilon^2}\sum_{m=1}^{\infty}E\left\{\max_{n_m\leqslant n<n_{m+1}}\left|\frac{1}{n}\sum_{1}^{n}x_j-\frac{1}{n}\sum_{1}^{n_m}x_j\right|^2\right\}<+\infty.$$

可见

$$\lim_{m\to\infty}\max_{n_m\leqslant n<n_{m+1}}\left|\frac{1}{n}\sum_{1}^{n}x_j-\frac{1}{n}\sum_{1}^{n_m}x_j\right|=0,\qquad \text{a.s.}. \quad (1.179)$$

对 $n_m\leqslant n<n_{m+1}$，则

$$\left|\frac{1}{n}\sum_{1}^{n}x_j\right|\leqslant\left|\frac{1}{n}\sum_{1}^{n}x_j-\frac{1}{n}\sum_{1}^{n_m}x_j\right|+\left|\frac{1}{n}\sum_{1}^{n_m}x_j-\frac{1}{n_m}\sum_{1}^{n_m}x_j\right|$$
$$+\left|\frac{1}{n_m}\sum_{1}^{n_m}x_j\right|, \quad (1.180)$$

其中第二项

$$\left|\frac{1}{n}\sum_{1}^{n_m}x_j-\frac{1}{n_m}\sum_{1}^{n_m}x_j\right|=\left|\frac{1}{n}-\frac{1}{n_m}\right|\left|\sum_{1}^{n_m}x_j\right|. \quad (1.181)$$

由于

$$\left|\frac{1}{n}-\frac{1}{n_m}\right|<\frac{1}{m^\beta}-\frac{1}{1+(m+1)^\beta}\leqslant C\frac{1}{m^\beta},\quad C\text{ 为常数},$$

故(1.181)为

$$\left|\frac{1}{n_m}-\frac{1}{n}\right|\left|\sum_1^{n_m}x_j\right|\leqslant C\frac{n_m}{m^\beta}\left|\frac{1}{n_m}\sum_1^{n_m}x_j\right|. \qquad (1.182)$$

利用(1.176)，(1.179)及(1.182)知

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_j=0,\qquad \text{a.s.}.$$

(2) 对于 $-\pi\leqslant\lambda\leqslant\pi$，$\lambda$ 固定，令

$$y_n=x_n e^{-in\lambda},\qquad n=0,\pm1,\cdots, \qquad (1.183)$$

则

$$R_y(k)=R_x(k)e^{-ik\lambda},\qquad k=0,\pm1,\cdots. \qquad (1.184)$$

因而（见(1.175)）

$$E\left\{\left|\frac{1}{n}\sum_{j=1}^{n}y_j\right|^2\right\}=E\left\{\left|\frac{1}{n}\sum_{j=1}^{n}x_je^{-ij\lambda}\right|^2\right\}$$

$$=\frac{1}{n}\sum_{l=-(n-1)}^{n-1}\left(1-\frac{|l|}{n}\right)R_x(l)e^{-il\lambda}\leqslant\frac{K}{n^a}, \qquad (1.185)$$

$K,a>0$ 为常数，引用前半部分的结果得，$-\pi\leqslant\lambda\leqslant\pi$ 有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}y_j=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_je^{-ij\lambda}=0,\qquad \text{a.s.}. \qquad (1.186)$$

定理全部得证。▎

推论1 若 x_t 的相关函数满足

$$\sum_k|R_x(k)|<+\infty,$$

则定理 1.16 的结论同样成立

证明 因为

$$\left|\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)R_x(j)e^{-ij\lambda}\right|\leqslant\frac{1}{n}\sum_{j=-\infty}^{\infty}|R_x(j)|\leqslant\frac{K}{n}$$

相当于(1.173)中 $\alpha=1$ 的场合，故结论仍成立。 ▎

推论2 若 x_t 的相关函数满足

$$|R_x(k)|<\frac{K}{n^\alpha},\quad \alpha>0, \tag{1.187}$$

则定理1.16结论仍成立。

证明 由(1.187)知

$$\left|\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)R_x(j)e^{-ij\lambda}\right|$$

$$\leqslant\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)|R_x(j)|\leqslant\frac{K}{n^\alpha}. \tag{1.188}$$

最后一不等式是因为

$$\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)\equiv 1. \tag{1.189}$$

从而定理结论仍正确。 ▎

推论3 设 x_t 为平稳列，$Ex_t\equiv a$，$R_x(k)$为其协方差函数，满足条件：

$$\frac{1}{n}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)R_x(j)=O\left(\frac{1}{n^\alpha}\right),\quad \alpha>0, \tag{1.190}$$

则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_j=Ex_t=a,\quad \text{a.s.}. \tag{1.191}$$

证明 令 $\tilde{x}_j=x_j-a$，$j=0,\pm1,\cdots$，则

$$R_{\tilde{x}}(n)=R_x(n).$$

故(1.190)可改为对 $R_{\tilde{x}}(j)$成立，因此

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}(x_j-a)=0,\quad \text{a.s.},$$

即(1.191)成立。 ▌

2° 关于相关函数的强大数律

利用定理1.16我们可证明以下在理论和应用上都十分重要的定理：

定理1.17 设 x_n 是平稳列(零均值)，并且 $X_n=x_{n+\nu}\bar{x}_n$ 亦平稳，即对固定的 ν，对一切 m

$$E\{(x_{\nu+n+m}\bar{x}_{n+m})(\overline{x_{n+\nu}\bar{x}_n})\}\tag{1.192}$$

与 n 无关，又若存在 K 和 $a>0$ 使得

$$\frac{1}{n}\sum_{j=-n-1}^{n-1}\left(1-\frac{|j|}{n}\right)R_X(j)-|R_x(\nu)|^2\leqslant\frac{K}{n^a},\tag{1.193}$$

则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_{\nu+j}\bar{x}_j=R_x(\nu),\quad \text{a.s.}.\tag{1.194}$$

证明 由 x_n 是平稳列，对固定的 ν，

$$EX_n=E\{x_{\nu+n}\bar{x}_n\}=R_x(\nu)$$

与 n 无关，则 $y_n=X_n-R_x(\nu)$ 是零均值平稳列，

$$R_y(j)=E(X_{n+j}\bar{X}_n)-|R_x(\nu)|^2,$$

$$\begin{aligned}&\frac{1}{n}\sum_{j=1-n}^{n-1}\left(1-\frac{|j|}{n}\right)R_y(j)\\&=\frac{1}{n}\sum_{j=1-n}^{n-1}\left(1-\frac{|j|}{n}\right)[E(X_{n+j}\bar{X}_n)-|R_x(\nu)|^2]\\&=\frac{1}{n}\sum_{j=1-n}^{n-1}\left(1-\frac{|j|}{n}\right)R_X(j)-|R_x(\nu)|^2\leqslant\frac{K}{n^a}\quad(a>0).\end{aligned}\tag{1.195}$$

利用定理1.16，则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}y_j=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}X_j-R_x(\nu)=0,$$

即

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_{j+\nu}\bar{x}_j=R_x(\nu),\quad \text{a.s.}. \quad ∎$$

推论 1 设 x_t 为零均值的平稳实 Gauss 序列，其相关函数满足

$$|R_x(n)|<\frac{K}{n^{\alpha}},\qquad K,\alpha>0, \tag{1.196}$$

则对任意的 ν，有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}x_{\nu-j}x_j=R_x(\nu),\quad \text{a.s.}.$$

证明 令 $y_n=X_n-R_x(\nu)$。则

$$R_y(n)=EX_{n-s}X_s-|R_x(\nu)|^2.$$

由 Gauss 性质①

$$\begin{aligned}EX_{n+s}X_s&=E(x_{\nu+n+s}x_{n+s}x_{s+\nu}x_s)\\&=R_x^2(\nu)+R_x^2(n)+R_x(n+\nu)R_x(n-\nu),\end{aligned} \tag{1.197}$$

从而

$$R_y(n)=R_x^2(n)+R_x(n+\nu)R_x(n-\nu),$$

$$|R_y(n)|\leqslant\frac{K^2}{n^{2\alpha}}+\frac{K^2}{(n+\nu)^{\alpha}}\cdot\frac{K^2}{(n-\nu)^{\alpha}}$$

$$=K^2\left(\frac{1}{n^{2\alpha}}+\frac{1}{(n^2-\nu^2)^{\alpha}}\right)<\frac{K_1}{n^{2\alpha}}\quad(2\alpha>0),$$

由定理1.16的推论 2

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}y_j=0,\qquad \text{a.s.}. \tag{1.198}$$

证毕。∎

事实上可以证明，比(1.196)更弱的条件下结论仍然成立。

定理1.18 设 x_t 为零均值平稳实 Gauss 序列，若其相关函

---

① 见本章习题12。

数满足：

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}|R_x(j)|^2=0, \tag{1.199}$$

或等价地：x_t 的谱函数 $F_x(\lambda)$在 Π 上是连续的，则(1.194)式仍正确。

此定理的证明可参看[1]，[2]。

由定理1.18立刻可得以下推论：

推论 在定理1.18条件下，若 x_t 的相关函数满足

$$\sum_{k=-\infty}^{\infty}|R_x(k)|<+\infty \tag{1.200}$$

或 x_t 具有谱密度 $f_x(\lambda)$，则定理1.18的结论仍成立。

3° 强大数定律的实际意义

以上我们证明的是：设 ξ_t 是平稳列，均值为 $E\xi_t$，在满足适当条件之下，它的样本序列(过程的一次实现) $\{x_j(\omega),\ j=1,2,\cdots,n\}$ 有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}x_k(\omega)=E\xi_t,\qquad \text{a.s.,}$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}x_{k+\tau}(\omega)x_k(\omega)=E\xi_{t+\tau}\xi_t=B(\tau),\qquad \text{a.s.,} \tag{1.201}$$

物理上的解释为：(1.201)的左边，如果把序列的足标看为时间，恰好代表了对过程(样本)的时间平均运算，而(1.201)右边是统计平均或称对总体的平均运算。也就是说平稳序列在一定的条件下具有性质：

$$\text{对时间求平均}=\text{对总体求平均}①. \tag{1.202}$$

---

① 在统计学中，以往学过的求 $E\xi_t$ 多指对ξ_t的多次样本观测 $\{x_t^{(i)},\ i=1,2,\cdots,N\}$ 求平均。

这一结果表明求过程的统计特征(矩)只需要过程的一次观测，而无需多个抽样的实现$\{x_n^{(i)}: n=1,2,\cdots,N; i=1,2,\cdots,M\}$。无疑，这对许多科学技术工作带来了极大的便利。工程技术上(或物理)称(1.202)的性质为"遍历的"或"各态历经的"或"Ergodic的"；数学上有时也称之为遍历性，但数学上通常讲的遍历性或Ergodic性质比(1.202)要广要深[1],[2]。

对连续参数的平稳过程在满足适当条件的情况下也有上述Ergodic性质。例如，比较常用的L^2连续的实平稳Gauss过程，有谱密度存在，则它是遍历的。设ξ_t是平稳过程，$\{x_t(\omega), 0\leqslant t\leqslant T\}$是它的一次实现样本，则

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x_t(\omega)\mathrm{d}t = E\xi_t, \quad \text{a.s.}. \qquad (1.203)$$

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x_{t+\tau}(\omega)x_t(\omega)\mathrm{d}t = B(\tau), \quad \text{a.s.}. \qquad (1.204)$$

习　题

1. 设$\xi_t=\sum_{k=1}^{N}A_k\cos(\omega_k t+\theta_k)$，$t\in\boldsymbol{R}_1$，$\{\theta_k\}$是独立同分布的随机变量，分布为$U(0,2\pi)$，$\{A_k\}$，$\{\omega_k\}$为常数。则$\xi_t$是平稳过程，试求其相关函数。

2. 设$\xi_t=\sum_{k=1}^{N}\xi_k e^{i\omega_k t}, t\in\boldsymbol{R}_1, \omega_k$不相等，$E|\xi_k|^2<+\infty$，$E\xi_k=0$，则$\xi_t$是平稳过程的充分必要条件是$E\xi_k\bar{\xi}_j=0(k\neq j)$。

3. 设$\xi_t=\xi f(t)$，$E|\xi|^2<+\infty$，$f(t)$是$\boldsymbol{R}_1$上的连续函数。则ξ_t是平稳过程的充分必要条件是$f(t)=r\,e^{i(\omega_0 t+\theta)}$，$r>0$，$\theta,\omega_0$为常数。

4. 设随机变量ξ有二阶矩，令$\xi_t\equiv\xi$，$t=0,\pm1,\pm2,\cdots$，

试证 ξ_t 是平稳列，并求其相关函数和谱测度 $F(\cdot)$，$\cdot\in\mathscr{B}(\Pi)$。

5. 设 $\xi_t, t\in R_1$ 是均方连续的实平稳过程，$E\xi_t=0$，其谱密度存在，即

$$f_\xi(\lambda)=\begin{cases} c\ (c>0), & \text{若}\ |\lambda|\leqslant 2\pi w; \\ 0, & \text{其它。}\end{cases}$$

则其相关函数 $R_\xi^{(w)}(\tau)$ 满足

$$\lim_{w\to\infty}\frac{R_\xi^{(w)}(\tau)}{R_\xi^{(w)}(0)}=\begin{cases}1, & \text{当}\ \tau=0; \\ 0, & \text{当}\ \tau\neq 0\text{。}\end{cases}$$

6. 设$\{Z(\lambda),\lambda\in\boldsymbol{R}_1\}$是 $\sigma=1$ 的 Wiener 过程，定义随机积分

$$\xi_t=\int_{\boldsymbol{R}_1} f(t-\lambda)\mathrm{d}Z(\lambda),\qquad t\in\boldsymbol{R}_1,$$

$f(t)\in L^2(\mathrm{d}\lambda)$。则$\{\xi_t,t\in\boldsymbol{R}_1\}$是平稳过程。试求其一维分布，如果取

A. $f(t)=\begin{cases} c\mathrm{e}^{-at}, & t\geqslant 0\ (c>0,a>0); \\ 0, & \text{其它。}\end{cases}$

B. $f(s)=\begin{cases} c\mathrm{e}^{-as}\sin\omega_0 s, & s\geqslant 0\ (c>0,\omega_0\neq 0,a>0); \\ 0, & s<0\text{。}\end{cases}$

试求 ξ_t 相应的相关函数和谱密度。

7. 设 $\xi_t=\sum\limits_{k=-\infty}^{\infty}a_k\eta_{t-k}$（在 H 内）$t=0,\pm 1,\cdots$，$E\eta_t=0$，$E\eta_t\eta_s=\delta_{t,s}$，则 ξ_t 必为平稳序列，其谱密度存在

$$f_\xi(\lambda)=\frac{1}{2\pi}\left|\sum_{k=-\infty}^{\infty}a_k\mathrm{e}^{-ik\lambda}\right|^2,\quad \lambda\in\Pi\text{。}$$

8. 设 $f(t)$是实连续函数，以 2π 为周期，并分段单调，ξ 是遵从 $U(-\pi,\pi)$分布的随机变量，则 $\xi_t=f(t+\xi)$，$t\in\boldsymbol{R}_1$ 是平稳过程。试求其谱展式，谱测度 $F(\cdot)$和正交测度 $Z(\cdot)$，$\cdot\in\mathscr{B}_1$。

9．设$Z(\cdot)$是$(\Lambda,\mathscr{B}(\Lambda),F)$上的随机正交测度，$F(\Lambda)<+\infty$，$f(t,\lambda)$是$T\times\Lambda$上的复值函数，对固定的$t\in T$，$f(t,\lambda)\in L^2(\mathrm{d}F)$，则

$$\xi_t=\int_\Lambda f(t,\lambda)\mathrm{d}Z(\lambda),\qquad t\in T$$

是二阶过程，试求其协方差函数。

10．设$\eta_t=a+bt+\xi_t$，$\{\xi_t\}$为独立同分布列，均值为0，方差为σ^2，a,b为非零常数。令

$$W_t=\frac{1}{2q+1}\sum_{j=-q}^{q}\eta_{t+j}.$$

试求其均值和协方差函数，并问它是否平稳？

11．设ε_t为独立同分布列，均值为0，方差为σ^2，a,b为非零常数，试问：以下哪些是平稳列？并求相关函数：

(1) $\xi_t=a+b\varepsilon_t+c\varepsilon_{t-1}$;

(2) $\xi_t=\varepsilon_t\cos ct+\varepsilon_{t-1}\sin ct$;

(3) $\xi_t=\varepsilon_t\varepsilon_{t-1}$;

(4) $\xi_t=\varepsilon_0\cos ct$.

12．设$\{\xi_t,t\in T\}$是零均值Gauss过程，则

$$E[\xi_{t_1}\xi_{t_2}\xi_{t_3}\xi_{t_4}]=E(\xi_{t_1},\xi_{t_2})E(\xi_{t_3}\xi_{t_4})+E(\xi_{t_1}\xi_{t_3})E(\xi_{t_2}\xi_{t_4})+E(\xi_{t_1}\xi_{t_4})E(\xi_{t_2}\xi_{t_3}).$$

又：设ξ_t的相关函数为$R(\tau)$，试求上式的表达式。

第二章　ARMA 模型

§1　ARMA模型与有理谱密度

1°　ARMA模型

在这一章里，我们将介绍一类具体的，在自然科学、工程技术以及社会、经济学的建模分析工作中起着非常重要作用的平稳随机序列模型，称为 ARMA 模型①，它的定义是：

定义2.1　设 ξ_t 是实平稳列($E\xi_t=0$)，它满足以下的随机差分方程

$$\xi_t+\varphi_1\xi_{t-1}+\cdots+\varphi_p\xi_{t-p}=\theta_0\varepsilon_t+\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}, \tag{2.1}$$

其中多项式

$$\Phi(z)=\sum_{k=0}^{p}\varphi_k z^k \quad (\varphi_0=1),$$

$$\Theta(z)=\sum_{k=0}^{q}\theta_k z^k \quad (\theta_0>0) \tag{2.2}$$

皆为实系数的多项式，它们的根都在 $|z|>1$ 的单位圆外，ε_t 是标准的白噪声序列，$E\varepsilon_t=0, E\varepsilon_t\varepsilon_s=\delta_{t,s}$，则称$\{\xi_t, t=0,\pm1,\cdots\}$为 ARMA$(p,q)$序列②，或称为 ARMA 模型。

如果$\Theta(z)$的根出现在$|z|\geqslant1$上，以后称为广 ARMA 模型。

若用U表推移算子(向后)，即

$$U^k\xi_t=\xi_{t-k}, \tag{2.3}$$

---

① ARMA 是AUTO REGRESSIVE MOVING AVERAGE 的缩写，意为“自迴归滑动平均”。

② p,q 通常指 ξ_t 能满足的 ARMA 方程的最小阶数。

则利用(2.2)，可将(2.1)简记为

$$\Phi(U)\xi_t = \Theta(U)\varepsilon_t. \tag{2.4}$$

定义2.2 在定义2.1中，若 $\theta_1 = \theta_2 = \cdots = \theta_q = 0$，则称满足方程

$$\xi_t + \varphi_1\xi_{t-1} + \cdots + \varphi_p\xi_{t-p} = \theta_0\varepsilon_t \tag{2.5}$$

的平稳列$\{\xi_t\}$为 AR(p)模型①。

定义2.3 在定义2.1中，若 $\varphi_1 = \varphi_2 = \cdots = \varphi_p = 0$，则称满足

$$\xi_t = \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} \tag{2.6}$$

的平稳列$\{\xi_t\}$为 MA(q)模型②。如果(2.6)的 $\Theta(z)$ 的根出现在 $|z|\geqslant 1$，则称为广 MA(q)模型。

例2.1 设 ξ_t 是方程

$$\xi_t - \frac{1}{2}\xi_{t-1} = 2\varepsilon_t \tag{2.7}$$

的平稳解，则它是AR(1)模型，因为

$$\Phi(z) = 1 - \frac{1}{2}z$$

只有一个$z=2$的根，$|z|>1$。

例2.2 设 ξ_t 满足

$$\xi_t = 6\varepsilon_t - 5\varepsilon_{t-1} + \varepsilon_{t-2},$$

则它是 MA(2)型，因为

$$\Theta(z) = 6 - 5z + z^2 = (3-z)(2-z)$$

有二个模大于 1 的实根，且 $\theta_0 = 6 > 0$。

例2.3 设 ξ_t 是满足方程

$$\xi_t - \frac{1}{2}\xi_{t-1} = 2\varepsilon_t - 3\varepsilon_{t-1} + \varepsilon_{t-2}$$

的平稳列，则 ξ_t 是广 ARMA(1,2)模型，因为 $\Theta(z)$ 根的模$\geqslant 1$。

---

① AR 是 AUTO REGRESSIVE(自回归)的缩写。

② MA 是 MOVING AVERAGE(滑动平均)的缩写。

例2.4 若ξ_t满足的方程为

$$\xi_t-\frac{3}{2}\xi_{t-1}+\frac{1}{2}\xi_{t-2}=\varepsilon_t-\frac{1}{2}\varepsilon_{t-1},$$

则它不是ARMA(2,1)模型，因为

$$\Phi(z)=\frac{1}{2}(z-1)(z-2)$$

有$z=1$的根。

在低阶条件下，ARMA模型中对$\Phi(z),\Theta(z)$根的判别可用以下的简单方法，一般高阶多项式根模的判别可参看第四章的附录或[10]。

A. ARMA(1,1)：设二个多项式为

$$\begin{cases}\Phi(z)=1+\varphi_1 z,\\ \Theta(z)=\theta_0+\theta_1 z,\end{cases}$$

第一个方程的根为$z_1=-\frac{1}{\varphi_1}$，因此要$|z_1|>1$只有$|\varphi_1|<1$。第二个方程的根为$z_2=-\theta_0/\theta_1$，因此$|z_2|>1$只有$|\theta_1/\theta_0|<1$。令$\lambda_1=\theta_1/\theta_0$，则ARMA(1,1)要求：

$$\begin{cases}|\lambda_1|<1,\\ |\varphi_1|<1.\end{cases}\tag{2.8}$$

B. AR(2)：设方程为

$$\xi_t+\varphi_1\xi_{t-1}+\varphi_2\xi_{t-2}=\theta_0\varepsilon_t,$$

对应的多项式为$\Phi(z)=1+\varphi_1 z+\varphi_2 z^2$，设它的两个根为$z_1$和$z_2$，显见应有

$$\begin{cases}z_1+z_2=-\varphi_1/\varphi_2,\\ z_1\cdot z_2=1/\varphi_2.\end{cases}\tag{2.9}$$

由实系数知z_1,z_2或同时为实数或为共轭数。若$|z_i|>1$，则由(2.9)第二式知$|\varphi_2|<1$。又由

$$-\varphi_1=\varphi_2(z_1+z_2)=\frac{z_1+z_2}{z_1z_2}=\frac{1}{z_1}+\frac{1}{z_2}$$

知

$$\left.\begin{aligned}\varphi_2-\varphi_1&=\frac{1}{z_1z_2}+\frac{1}{z_1}+\frac{1}{z_2}=\left(1+\frac{1}{z_1}\right)\left(1+\frac{1}{z_2}\right)-1,\\ \varphi_2+\varphi_1&=\frac{1}{z_1z_2}-\left(\frac{1}{z_1}+\frac{1}{z_2}\right)=\left(1-\frac{1}{z_1}\right)\left(1-\frac{1}{z_2}\right)-1.\end{aligned}\right\}\tag{2.10}$$

以下说明：无论 z_1,z_2 是实数或共轭复数，都有

$$\left(1+\frac{1}{z_1}\right)\left(1+\frac{1}{z_2}\right)\tag{2.11}$$

和

$$\left(1-\frac{1}{z_1}\right)\left(1-\frac{1}{z_2}\right)\tag{2.12}$$

皆大于零。事实上，若 z_i 是模大于 1 的实数，结论显然成立。若 z_i 为共轭复数，则(2.12)为

$$\frac{(1-z_1)(1-\bar{z}_1)}{z_1\bar{z}_1}=\left|\frac{1-z_1}{z_1}\right|^2=\frac{|1-z_1|^2}{|z_1|^2}>0.$$

否则表明 $z_1=1$，和模大于 1 的要求矛盾。同理可证

$$\frac{(1+z_1)(1+\bar{z}_1)}{z_1\bar{z}_1}>0.$$

由以上事实及(2.10)式知

$$\begin{cases}-1<\varphi_2-\varphi_1,\\ -1<\varphi_2+\varphi_1.\end{cases}\tag{2.13}$$

因此AR(2)的条件是

$$\begin{cases}|\varphi_2|<1,\\ \varphi_1-\varphi_2<1,\\ -1<\varphi_1+\varphi_2.\end{cases}\tag{2.14}$$

反之，读者可以证明，若 $\Phi(z)$ 的系数满足(2.14)，则它的两个

根必满足$|z_i|>1(i=1,2)$。

C. MA(2)：设多项式为

$$\Theta(z)=\theta_0+\theta_1 z+\theta_2 z^2.$$

Θ可变形为：

$$\Theta(z)=\theta_0(1+\lambda_1 z+\lambda_2 z^2),$$

其中

$$\lambda_1=\theta_1/\theta_0,\quad \lambda_2=\theta_2/\theta_0. \tag{2.15}$$

由AR(2)的判别法知MA(2)的条件为

$$\begin{cases}|\lambda_2|<1,\\ \lambda_1-\lambda_2<1,\\ -1<\lambda_1+\lambda_2.\end{cases} \tag{2.16}$$

2° ARMA 模型的平稳解

为了得到一般ARMA模型平稳解的解析表达式，我们先来看一个实例。

设ξ_t是方程

$$\xi_t+\varphi_1\xi_{t-1}=\theta_0\varepsilon_t,\qquad |\varphi_1|<1,\ \theta_0>0 \tag{2.17}$$

的平稳解。将(2.17)写成递推形式可得

$$\begin{aligned}\xi_t&=-\varphi_1\xi_{t-1}+\theta_0\varepsilon_t\\ &=-\varphi_1[-\varphi_1\xi_{t-2}+\theta_0\varepsilon_{t-1}]+\theta_0\varepsilon_t\\ &=\sum_{k=0}^{n}\theta_0(-1)^k\varphi_1^k\varepsilon_{t-k}+(-1)^{n+1}\varphi_1^{n+1}\xi_{t-(n+1)}\\ &=\sum_{k=0}^{n}c_k\varepsilon_{t-k}+(-\varphi_1)^{n+1}\xi_{t-(n+1)},\end{aligned} \tag{2.18}$$

其中

$$c_k=\theta_0(-\varphi_1)^k.$$

$$\begin{aligned}\left\|\xi_t-\sum_{k=0}^{n}c_k\varepsilon_{t-k}\right\|^2&=\|(-\varphi_1)^{n+1}\xi_{t-(n+1)}\|^2\\ &=|\varphi_1|^{2(n+1)}\|\xi_0\|^2\to 0\quad(n\to\infty).\end{aligned} \tag{2.19}$$

可见在H中ξ_t可表为L^2级数形式

$$\xi_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}, \tag{2.20}$$

即ξ_t可表为白噪声列$\{\varepsilon_t\}$的半边滑动和（MA($+\infty$)）。其次，若取$\beta = -\ln|\varphi_1|>0$，则

$$|c_k| = |\theta_0(-\varphi_1)^k| \leqslant \theta_0 |\varphi_1|^k = \theta_0 e^{-\beta k}, \tag{2.21}$$

$$k = 0,1,2,\cdots.$$

这表明(2.20)的系数是指数衰减的。将ε_t的谱展式

$$\varepsilon_t = \int_{\Pi} e^{i\lambda t} dZ_\varepsilon(\lambda)$$

代入(2.20)可得

$$\begin{aligned}
\xi_t &= \int_{\Pi} e^{i\lambda t}\left(\sum_{k=0}^{\infty} c_k e^{-ik\lambda}\right) dZ_\varepsilon(\lambda) \\
&= \int_{\Pi} e^{i\lambda t}\left(\theta_0 \sum_{k=0}^{\infty} (-\varphi_1)^k e^{-ik\lambda}\right) dZ_\varepsilon(\lambda) \\
&= \int_{\Pi} e^{i\lambda t} \frac{\theta_0}{1-(-\varphi_1)e^{-i\lambda}} dZ_\varepsilon(\lambda) \\
&= \int_{\Pi} e^{i\lambda t} \frac{\theta_0}{\Phi(e^{-i\lambda})} dZ_\varepsilon(\lambda).
\end{aligned} \tag{2.22}$$

其中$\Phi(e^{-i\lambda}) = \Phi(z)|_{z=e^{-i\lambda}} = 1 + \varphi_1 e^{-i\lambda}$。对照(2.17)可知$\xi_t$的表达式可用方程的系数表出。

(2.20)，(2.21)和(2.22)是研究ARMA模型的三条重要性质。为了论证一般ARMA模型也有这些结果需要先证明以下引理：

引理2.1 设$Y(\cdot)$是$(\Pi, \mathscr{B}(\Pi), F_Y)$上的正交随机测度$(F_Y(\Pi) < +\infty)$。如果$g \in L^2(dF_Y)$，则

$$Z(s) = \int_s g(\lambda) dY(\lambda), \quad s \in \mathscr{B}(\Pi) \tag{2.23}$$

是定义在$(\Pi, \mathscr{B}(\Pi), F_Z)$上的正交随机测度，其中

$$F_Z(A) = \int_A |g(\lambda)|^2 \mathrm{d}F_Y, \qquad A \in \mathscr{B}(\Pi). \tag{2.24}$$

证明 由正交随机测度的定义1.9，在H中

$$\begin{aligned}(Z(s_1), Z(s_2)) &= \left(\int_{s_1} g\mathrm{d}Y, \int_{s_2} g\mathrm{d}Y\right)_H \\ &= \int_\Pi (\chi_{s_1} g)(\overline{\chi_{s_2} g}) \mathrm{d}F_Y \\ &= \int_{s_1 \cap s_2} |g|^2 \mathrm{d}F_Y \\ &= F_Z(s_1 \cap s_2), \qquad s_1, s_2 \in \mathscr{B}(\Pi).\end{aligned} \tag{2.25}$$

可见$Z(\cdot)$是正交测度，

$$F_Z(\Pi) = \int_\Pi |g(\lambda)|^2 \mathrm{d}F_Y < +\infty. \quad ▌$$

引理2.2 在引理2.1条件下，若$f \in L^2(\mathrm{d}F_Z)$，则

$$\int_\Pi f(\lambda) \mathrm{d}Z(\lambda) = \int_\Pi f(\lambda) g(\lambda) \mathrm{d}Y(\lambda). \tag{2.26}$$

证明 设$f = \sum_1^n a_k \chi_{A_k} \in L\{\chi_{A_i}\}$，则

$$\begin{aligned}\int_\Pi f(\lambda) \mathrm{d}Z(\lambda) &= \sum_1^n a_k \int_{A_k} \mathrm{d}Z(\lambda) \\ &= \sum_1^n a_k \int_\Pi \chi_{A_k} g\mathrm{d}Y \qquad (\text{由}(2.23)) \\ &= \int_\Pi \left(\sum_1^n a_k \chi_{A_k}\right) g\mathrm{d}Y = \int_\Pi fg\mathrm{d}Y.\end{aligned} \tag{2.27}$$

表明(2.26)成立。

对$f\in L^2(\mathrm{d}F_Z)$，存在一串$f_n\in L\{\chi_{A_t}\}$，使

$$f_n\xrightarrow{L^2(\mathrm{d}F_Z)}f.$$

$$\begin{aligned}\|f-f_n\|^2_{L^2(\mathrm{d}F_Z)}&=\int_\Pi|f-f_n|^2\mathrm{d}F_Z\\&=\int_\Pi|f-f_n|^2|g|^2\mathrm{d}F_Y\\&=\int_\Pi|fg-f_ng|^2\mathrm{d}F_Y.\end{aligned}\qquad(2.28)$$

然而(2.28)的左端→0 ($n\to\infty$)，可见

$$f_ng\xrightarrow{L^2(\mathrm{d}F_Y)}fg,$$

即

$$\|f_ng-fg\|^2_{L^2(\mathrm{d}F_Y)}=\left\|\int_\Pi f_ng\mathrm{d}Y-\int_\Pi fg\mathrm{d}Y\right\|^2_H\to0.$$

表明

$$\int_\Pi fg\mathrm{d}Y=\lim_n\int_\Pi f_ng\mathrm{d}Y\overset{1}{=}\lim_n\int_\Pi f_n\mathrm{d}Z\overset{2}{=}\int_\Pi f\mathrm{d}Z.$$

最后两步的1是因为(2.27)已证，2是利用随机积分的定义。▮

定理2.1 设ξ_t是ARMA(p,q)

$$\Phi(U)\xi_t=\Theta(U)\varepsilon_t$$

的平稳解，则ξ_t可唯一地表为

$$\xi_t=\int_\Pi e^{i\lambda t}\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\mathrm{d}Z_\varepsilon(\lambda),\qquad(2.29)$$

其中$\mathrm{d}Z_\varepsilon$是平稳列ε_t的正交随机测度。

证明 首先验ξ_t是解。设

$$\Phi(z)=\sum_{k=0}^p\varphi_kz^k.$$

则

$$\begin{aligned}
\Phi(U)\xi_t &= \sum_{k=0}^{p}\varphi_k U^k \xi_t \\
&= \sum_{k=0}^{p}\varphi_k \int_{\Pi} e^{i\lambda(t-k)} \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} dZ_\varepsilon(\lambda) \\
&= \int_{\Pi} e^{i\lambda t} \sum_{k=0}^{p}\varphi_k e^{-ik\lambda} \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} dZ_\varepsilon(\lambda) \\
&= \int_{\Pi} e^{i\lambda t}\Theta(e^{-i\lambda}) dZ_\varepsilon(\lambda) \\
&= \int_{\Pi}\left(\sum_{k=0}^{q}\theta_k e^{i\lambda(t-k)}\right) dZ_\varepsilon(\lambda) \\
&= \Theta(U)\varepsilon_t.
\end{aligned} \tag{2.30}$$

由于 $\Theta(e^{-i\lambda})/\Phi(e^{-i\lambda})$是 Π 上的连续函数，Φ，Θ 多项式的根不出现在单位圆上，故存在常数$M,m>0$，使得

$$0<m<\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|<M, \qquad \lambda\in\Pi. \tag{2.31}$$

因此(2.29)的随机积分是存在的，并且

$$\begin{aligned}
R_\xi(t,\tau) &= \left(\int_{\Pi} e^{i\lambda(t+\tau)} \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} dZ_\varepsilon, \int_{\Pi} e^{i\lambda t} \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} dZ_\varepsilon\right)_H \\
&= \int_{\Pi} e^{i\lambda\tau} \left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2 dF_\varepsilon \\
&= \frac{1}{2\pi}\int_{\Pi} e^{i\lambda\tau} \left|\Theta(e^{-i\lambda})/\Phi(e^{-i\lambda})\right|^2 d\lambda \\
&= R_\xi(\tau).
\end{aligned} \tag{2.32}$$

表明 ξ_t 是平稳的，它有谱密度

$$f_\xi(\lambda) = \frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2, \qquad \lambda\in\Pi. \tag{2.33}$$

以下证 ξ_t 表达式(2.29)的唯一性。

设 ARMA(p,q)有另外的平稳解

$$\eta_t=\int_{\Pi} e^{i\lambda t}\,dZ_\eta(\lambda). \tag{2.34}$$

由于满足 ARMA 方程必有

$$\begin{aligned}\zeta_t\triangleq \Phi(U)\eta_t&=\int_{\Pi} e^{i\lambda t}\,\Phi(e^{-i\lambda})dZ_\eta\\ &=\Theta(U)\varepsilon_t=\Phi(U)\xi_t=\int_{\Pi} e^{i\lambda t}\Phi(e^{-i\lambda})dZ_\xi. \end{aligned}\tag{2.35}$$

然而$\zeta_t=\Theta(U)\varepsilon_t$是平稳列，由谱展式定理必可表为

$$\zeta_t=\int_{\Pi} e^{i\lambda t}\,dZ_\zeta(\lambda). \tag{2.36}$$

令

$$Z_1(A)=\int_A \Phi(e^{-i\lambda})dZ_\xi(\lambda)$$

$$Z_2(A)=\int_A \Phi(e^{-i\lambda})dZ_\eta(\lambda),\quad A\in\mathscr{B}(\Pi).$$

由(2.31)，$Z_i(\cdot)$皆有定义(引理2.1)。

$$\begin{aligned}\int_{\Pi} e^{i\lambda t}\,dZ_1(\lambda)&=\int_A e^{i\lambda t}\,\Phi(e^{-i\lambda})dZ_\xi(\lambda),\\ \int_{\Pi} e^{i\lambda t}\,dZ_2(\lambda)&=\int_A e^{i\lambda t}\,\Phi(e^{-i\lambda})dZ_\eta(\lambda).\end{aligned}\tag{2.37}$$

由(2.35)，(2.36)，(2.37)得：

$$\int_{\Pi} e^{i\lambda t}\,dZ_\zeta(\lambda)=\int_{\Pi} e^{i\lambda t}\,dZ_1(\lambda)=\int_{\Pi} e^{i\lambda t}\,dZ_2(\lambda). \tag{2.38}$$

由ζ_t谱展式的唯一性，

$$Z_\zeta(A)=Z_1(A)=Z_2(A),\qquad A\in\mathscr{B}(\Pi). \tag{2.39}$$

由于$(\Phi(e^{-i\lambda}))^{-1}$有界，可考虑随机积分

$$\int_A \frac{dZ_\zeta(\lambda)}{\Phi(e^{-i\lambda})}\overset{*}{=}\int_A \frac{\Phi(e^{-i\lambda})}{\Phi(e^{-i\lambda})}dZ_\xi(\lambda)=Z_\xi(A),\quad A\in\mathscr{B}(\Pi),$$

其中$*$是由引理2.2作测度变换。同理对$Z_\eta(\cdot)$亦有

$$\int_A \frac{dZ_\zeta(\lambda)}{\Phi(e^{-i\lambda})}=\int_A \frac{\Phi(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\,dZ_\eta(\lambda)=Z_\eta(A),\qquad A\in\mathscr{B}(\Pi).$$

可见 $Z_\xi(A)=Z_\eta(A)$，$A\in\mathscr{B}(\Pi)$，因而 $\eta_t=\xi_t$。 ∎

以下定理表明 ARMA 序列 ξ_t 必可表成 $\{\varepsilon_t\}$ 的单边滑动和。

定理2.2 设 ξ_t 是 ARMA(p,q)

$$\Phi(U)\xi_t=\Theta(U)\varepsilon_t$$

的平稳解，则 ξ_t 可唯一地表为

$$\xi_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k}\ (H\text{内}),\tag{2.40}$$

其中 $\{c_k,k\geqslant 0\}$ 是解析函数

$$\Gamma_\xi(z)=\frac{\Theta(z)}{\Phi(z)}\tag{2.41}$$

在单位圆内的 Taylor 系数，$\{c_k\}\in l_1$。

证明 由 $\Phi(z),\Theta(z)$ 的假定，可得

$$\Phi(z)=\prod_{j=1}^{p}(z-\beta_j),\qquad |\beta_j|>1;$$

$$\Theta(z)=\theta_0\prod_{l=1}^{q}(z-\alpha_l),\qquad |\alpha_l|>1.$$

$$\frac{\Theta(z)}{\Phi(z)}=\theta_0\prod_{l=1}^{q}(z-\alpha_l)\prod_{j=1}^{p}(z-\beta_j)^{-1}$$

是单位圆内(可比单位圆大一点)的解析函数，因而可 Taylor 展开得

$$\frac{\Theta(z)}{\Phi(z)}=\sum_{k=0}^{\infty}c_kz^k,\tag{2.42}$$

它在 $|z|\leqslant 1$ 内绝对收敛，因而对 $z=e^{-i\lambda}$ 有

$$\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}=\sum_{k=0}^{\infty}c_ke^{-ik\lambda},\tag{2.43}$$

并且 $\{c_k\}$ 是绝对收敛的，即 $\{c_k\}\in l_1$。从而由定理2.1知

$$\xi_t=\int_{\Pi} e^{i\lambda t}\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}dZ_{\varepsilon}(\lambda)$$

$$=\sum_{k=0}^{\infty}c_k\int_{\Pi}e^{i\lambda(t-k)}dZ_{\varepsilon}(\lambda)$$

$$=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k}. \tag{2.44}$$

$\{c_k\}$的唯一性是由(2.43)Taylor 展开的唯一性得来的。 ▎

(2.41)式所定义的 $\Gamma_{\xi}(z)$在时间序列的研究中起着重要的作用，它满足$\Gamma_{\xi}(z)\neq 0$，$|z|\leqslant 1$，以后我们称它为ξ_t的 ARMA 方程所对应的极大解析函数，或简称为ξ_t的极大解析函数。由(2.33)，ARMA 序列ξ_t的谱密度$f_{\xi}(\lambda)$可表为

$$f_{\xi}(\lambda)=\frac{1}{2\pi}|\Gamma_{\xi}(e^{-i\lambda})|^2,\quad \lambda\in\Pi. \tag{2.45}$$

以后称(2.40)的展式为 Wold 展式，$\{c_k,k\geqslant 0\}$称为ξ_t的 Wold 系数。

由于在定理2.1，2.2的证明中$\{\varepsilon_t\}$白噪声的性质并没有起特别的作用，因此将 ARMA 方程改写为

$$\Theta(U)\varepsilon_t=\Phi(U)\xi_t, \tag{2.46}$$

则不难证得如下的定理:

定理2.3 设ξ_t是 ARMA 方程

$$\Phi(U)\xi_t=\Theta(U)\varepsilon_t$$

的平稳解，则

$$\varepsilon_t=\int_{\Pi}e^{i\lambda t}\Gamma_{\xi}^{-1}(e^{-i\lambda})dZ_{\xi}(\lambda) \tag{2.47}$$

$$=\sum_{k=0}^{\infty}d_k\xi_{t-k}, \tag{2.48}$$

其中$\{d_k\}\in l_1$，是

$$\Gamma_{\xi}^{-1}(z)=\frac{\Phi(z)}{\Theta(z)} \tag{2.49}$$

在单位圆内展开的 Taylor 系数。

3° ARMA 模型的谱密度

在(2.45)式中已表示出一般 ARMA 模型的谱密度，事实上这一结果也是充分的：

定理2.4 设 $\Phi(z),\Theta(z)$ 为两个 p 阶和 q 阶的多项式，满足定义2.1中对 Φ,Θ 的条件，ξ_t 是实平稳列。则 ξ_t 是 ARMA(p,q) 方程

$$\Phi(U)\xi_t=\Theta(U)\varepsilon_t$$

的解的充分必要条件是 $F_\xi(\lambda)$ 谱函数绝对连续，并且谱密度为

$$f_\xi(\lambda)=\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2,\quad \lambda\in\Pi \tag{2.50}$$

证明 必要性已见于(2.45)式。下面证充分性。

由于 ξ_t 是平稳列，由谱展式，

$$\xi_t=\int_\Pi e^{i\lambda t}\,dZ_\xi(\lambda).$$

令

$$\varepsilon_t=\int_\Pi e^{i\lambda t}\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}dZ_\xi(\lambda), \tag{2.51}$$

则

$$\begin{aligned}\sum_{k=0}^{q}\theta_k\varepsilon_{t-k}&=\int_\Pi\left(\sum_{k=0}^{q}\theta_k e^{i\lambda(t-k)}\right)\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}dZ_\xi(\lambda)\\&=\int_\Pi e^{i\lambda t}\left(\sum_{l=0}^{p}\varphi_l e^{-il\lambda}\right)dZ_\xi(\lambda)\\&=\sum_{l=0}^{p}\varphi_l\xi_{t-l},\end{aligned}$$

即 $\Phi(U)\xi_t=\Theta(U)\varepsilon_t$ 成立，其中$\{\varepsilon_t\}$为白噪声列：

$$(\varepsilon_{t+\tau},\varepsilon_t)_H$$
$$=\left(\int_\Pi e^{i\lambda(t+\tau)}\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}dZ_\xi(\lambda),\int_\Pi e^{i\lambda t}\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}dZ_\xi(\lambda)\right)_H$$

$$=\int_{\Pi} e^{i\lambda\tau}\left|\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}\right|^2 dF_{\xi}(\lambda)$$

$$=\frac{1}{2\pi}\int_{\Pi} e^{i\lambda\tau}\left|\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}\right|^2\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2 d\lambda \quad (由(2.50))$$

$$=\frac{1}{2\pi}\int_{\Pi} e^{i\lambda\tau}\, d\lambda=\delta_{0,\tau}. \tag{2.52}$$

这表明$\{\varepsilon_t\}$是白噪声列。

推论1 设$\Phi(z)$,$\Theta(z)$是满足定理2.4 条件的两个多项式，则必存在一个 ARMA(p,q)序列$\{\xi_t\}$ 以 $\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2$为谱密度。

证明 令

$$R(\tau)=\int_{\Pi} e^{i\lambda\tau}\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2 d\lambda, \quad \tau=0,\pm 1,\cdots, \tag{2.53}$$

则对任意的复数$(a_1,a_2,\cdots,a_n)$和整数$(\tau_1,\tau_2,\cdots,\tau_n)$,

$$\sum_{k=1}^{n}\sum_{l=1}^{n} a_k\bar{a}_l R(\tau_k-\tau_l)$$

$$=\int_{\Pi}\left(\sum_{k=1}^{n}\sum_{l=1}^{n} a_k\bar{a}_l e^{i(\tau_k-\tau_l)\lambda}\right)\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2 d\lambda$$

$$=\int_{\Pi}\left|\sum_{k=1}^{n} a_k e^{i\tau_k\lambda}\right|^2\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2 d\lambda\geqslant 0. \tag{2.54}$$

这表明 $R(\tau)$是非负定函数。由第一章中的 Gauss 存在定理（见定理1.2）取期望为零，协方差函数为

$$\sigma_{t,s}=R(t-s), \quad t,s \text{ 整数},$$

则必存在 Gauss 平稳列 ξ_t 以 $R(\tau)$为相关函数，由(2.53) 及谱密度的唯一性，必有

$$f_{\xi}(\lambda)=\frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2, \quad \lambda\in\Pi. \quad \blacksquare$$

推论2 ξ_t 是 AR(p)方程

$$\Phi(U)\xi_t=\theta_0\varepsilon_t$$

的平稳解的充分必要条件是 ξ_t 的谱密度存在，且

$$f_\xi(\lambda)=\frac{\theta_0^2}{2\pi|\Phi(e^{-i\lambda})|^2}, \qquad \lambda\in\Pi. \tag{2.55}$$

推论3 ξ_t 是 MA(q)序列

$$\xi_t=\Theta(U)\varepsilon_t$$

的充分必要条件是 ξ_t 谱密度存在，且

$$f_\xi(\lambda)=\frac{1}{2\pi}|\Theta(e^{-i\lambda})|^2, \qquad \lambda\in\Pi. \tag{2.56}$$

§2 ARMA 模型下 H_ξ 空间中的标准正交基

1° H_ξ 空间中的标准正交基

在第一章中已引入了概率空间 $(\Omega,\mathscr{F},P)$上具有二阶矩的全体随机变量的 Hilbert 空间H。假若$\{\xi_t,t=0,\pm1,\cdots\}$是平稳列，我们可以讨论比$H$更具体的子空间 H_ξ，它是由$\{\xi_t\}$张成的线性闭包

$$H_\xi=\mathscr{L}\{\xi_t;t=0,\pm1,\cdots\}. \tag{2.57}$$

假若$\{\xi_t\}$又是 ARMA 模型，由定理2.2的(2.40)式知

$$\xi_t\in H_\varepsilon=\mathscr{L}\{\varepsilon_{t'};t'=0,\pm1,\cdots\}. \tag{2.58}$$

又由定理2.3的(2.48)式知

$$\varepsilon_t\in H_\xi. \tag{2.59}$$

由(2.58)，(2.59)知必有

$$H_\xi=H_\varepsilon. \tag{2.60}$$

然而$\{\xi_t,t=0,\pm1,\cdots\}$在 H_ξ 中虽是生成元，但却不是正交系，因为

$$(\xi_t,\xi_s)=R(t-s),$$

当 $t\neq s$ 时并不恒为零。但是由(2.60)，$\{\varepsilon_t,t=0,\pm1,\cdots\}$不仅是生成元而且在 H_ξ 中还是正交系

$$(\varepsilon_t, \varepsilon_s) = 0, \qquad t \neq s.$$

定理2.5 设ξ_t是 ARMA 方程

$$\Phi(U)\xi_t = \Theta(U)\varepsilon_t$$

的平稳解，则

A．$\varepsilon_t \in H_\xi(t)$，

B．$\varepsilon_t \in H_\xi(t) \ominus H_\xi(t-1)$。

C．$\{\varepsilon_t, t=0, \pm 1, \cdots\}$是$H_\xi$中的一组完全标准正交基。

D．$c_0 > 0$。

证明

A．由$\varepsilon_t = \sum\limits_{k=0}^{\infty} d_k \xi_{t-k}$（见(2.48)式）及$H_\xi(t)$的定义

$$H_\xi(t) = \mathscr{L}\{\xi_k; k \leqslant t\}$$

知结论A成立。

B．对任意的$k>0$，由(2.40)式，

$$(\xi_{t-k}, \varepsilon_t) = \left(\sum_{l=0}^{\infty} c_l\, \varepsilon_{t-k-l}, \varepsilon_t\right)$$

$$= \sum_{l=0}^{\infty} c_l (\varepsilon_{t-k-l}, \varepsilon_t).$$

但$t-k-l<t$（因为$k>0$，$l \geqslant 0$），由正交性，

$$(\xi_{t-k}, \varepsilon_t) \equiv 0, \qquad k>0. \tag{2.61}$$

可见

$$\varepsilon_t \perp H_\xi(t-k), \qquad k>0. \tag{2.62}$$

结合A的结论，B成立。

C．标准正交性已证过，只需证完全性。设有$\zeta \in H_\xi$，且

$$(\zeta, \varepsilon_t) \equiv 0, \qquad \forall\, t. \tag{2.63}$$

由(2.60)式知ζ必可表为

$$\zeta = \lim_{n\to\infty} \sum_{k=0}^{n} c_k^{(n)}\, \varepsilon_{t_k^{(n)}}. \tag{2.64}$$

$$(\zeta,\zeta)=\lim_{n\to\infty}\sum_{k=0}^{n}c_k^{(n)}(\varepsilon_{t_k^{(n)}},\zeta)=0。\qquad(2.65)$$

可见 $\zeta=0$，表明$\{\varepsilon_t\}$是完全的。

D. 由$\{c_k\}$之定义（见定理2.2），

$$c_0=\Gamma_\xi(0)=\frac{\theta_0+\theta_1 z+\cdots+\theta_q z^q}{1+\varphi_1 z+\cdots+\varphi_p z^p}\bigg|_{z=0}=\theta_0>0。\qquad(2.66)$$

定理证毕。▌

推论 在定理2.5条件下，ξ_t 的Wold 系数

$$c_k=(\xi_t,\varepsilon_{t-k}),\qquad k\geqslant 0\qquad(2.67)$$

乃是 H_ξ 中 ξ_t 对正交基$\{\varepsilon_t\}$的广义 Fourier 系数。

证明 由 ARMA 模型的 Wold 展式

$$\xi_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k}$$

及$\{\varepsilon_t\}$的正交性，对 $k\geqslant 0$，

$$(\xi_t,\varepsilon_{t-k})=\sum_{s=0}^{\infty}c_s(\varepsilon_{t-s},\varepsilon_{t-k})=c_k。\ ▌$$

由上可见，ARMA 模型 ξ_t 的 Wold 展式

$$\xi_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k}\qquad(2.68)$$

正是 ξ_t 按$\{\varepsilon_t\}$展开的广义 Fourier 级数。

值得注意的是：在第一章定理1.14中，我们指出：平稳列$\{\xi_t\}$若有谱密度，则必可展为

$$\xi_t=\sum_{k=-\infty}^{\infty}a_k e_{t-k}。\qquad(2.69)$$

但是，此级数对$\{e_t\}$正交系的展开是双边的求和（从 $-\infty$ 到 $+\infty$），而 Wold 展式是单边求和。此外，ξ_t 展式的正交系$\{e_t\}$未必具有定理2.5的A，B，D性质，而这些性质在随机过程的预测、滤波理论中起着重要作用。

定理2.5在比 ARMA 模型更广的平稳列的条件下仍可成立。Колмогоров 在[13]中证明了：若平稳列$\{\xi_t\}$的谱函数 $F_\xi(\lambda)$满足

(i) $dF_\xi \ll d\lambda$ （即 dF_ξ 对 Lebegue 测度绝连续），

(ii) $\ln f_\xi(\lambda) \in L_1(d\lambda)$，即

$$\int_{\Pi} \ln f_\xi(\lambda) d\lambda > -\infty, \tag{2.70}$$

则 ξ_t 必有(2.68)的Wold 展式，而$\{c_k\}$也是由(2.67)确定，$\{\varepsilon_t\}$也具有定理2.5的 A,B,C,D 性质。

定理2.6 设 ξ_t 为 ARMA 序列。记①

$$E[\xi_t | H_\xi(t-\tau)] \triangleq \mathop{\mathrm{Proj}}_{H_\xi(t-\tau)} \{\xi_t\}, \quad \tau \geqslant 0. \tag{2.71}$$

则 ARMA 方程中的$\{\varepsilon_t\}$为

$$\varepsilon_t = \frac{\xi_t - E[\xi_t | H_\xi(t-1)]}{\|\xi_t - E[\xi_t | H_\xi(t-1)\|^2}, \tag{2.72}$$

其中 $t = \pm 1, \pm 2, \cdots$

证明 由Wold 展式，

$$\xi_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}.$$

由 Hilbert 空间中投影算子的线性性质可得

$$E[\xi_t | H_\xi(t-1)] = \sum_{k=0}^{\infty} c_k E[\varepsilon_{t-k} | H_\xi(t-1)].$$

由于
$$\varepsilon_{t-k} \in H_\xi(t-1), \quad k \geqslant 1,$$
$$\varepsilon_t \perp H_\xi(t-1) \text{ （见(2.62)式），}$$

故

$$E[\xi_t | H_\xi(t-1)] = \sum_{k=1}^{\infty} c_k \varepsilon_{t-k}, \tag{2.73}$$

并且

$$\xi_t - E[\xi_t | H_\xi(t-1)] = c_0 \varepsilon_t. \tag{2.74}$$

从而

---

① $\mathop{\mathrm{Proj}}_{\mathscr{A}}(\xi)$表示$\xi$在了空间$\mathscr{A}$上的投影。

$$\|\xi_t - E[\xi_t | H_\xi(t-1)]\|^2 = c_0^2 > 0 \tag{2.75}$$

$$\varepsilon_t = \frac{\xi_t - E[\xi_t | H_\xi(t-1)]}{\|\xi_t - E[\xi_t | H_\xi(t-1)]\|}. \quad \blacksquare$$

由(2.72)可看出：ε_t 是用 $H_\xi(t-1)$ 里的元素所不能"认识"的 ξ_t 的"成份"，因为用$\{\xi_s, s \leqslant t-1\}$的一切可能线性组合来预测 ξ_t 的最小误差至少是 c_0(见(2.75))。在近代控制理论中称$\{\varepsilon_t\}$序列为"新息"序列 (Innovation Series)，在[13]中称为 Wold 序列。由(2.72)式可看出 Wold 展式、新息序列等概念可以推广到非平稳的场合中去[18]。

c_0 是平稳随机序列的一步预测误差。如果 $c_0 > 0$，则表明过程永远不可能"无误差"预测，这种过程我们称为是非奇异的。反之，如果 $c_0 = 0$，我们称该过程是奇异的。由定理2.5的D知，ARMA 序列是非奇异的。

2° ARMA 序列求 Wold 系数的递推公式

在上一节中，我们已初步看出 Wold 分解式(2.68)的重要性。求 Wold 系数$\{c_k, k \geqslant 0\}$虽可由原始定义对 $\Gamma_\xi(z)$ 在 $z=0$ 求 Taylor 系数，但实用中下面的递推公式更简易。

$$\Gamma_\xi(z) = \frac{\Theta(z)}{\Phi(z)} = \sum_{k=0}^{\infty} c_k z^k, \qquad |z| \leqslant 1,$$

$$\sum_{k=0}^{q} \theta_k z^k = \sum_{k=0}^{\infty} c_k \sum_{j=0}^{p} \varphi_j z^{k+j}. \tag{2.76}$$

令

$$\tilde{\theta}_k = \begin{cases} \theta_k, & 0 \leqslant k \leqslant q; \\ 0, & \text{其它。} \end{cases} \tag{2.77}$$

$$\tilde{\varphi}_s = \begin{cases} \varphi_s, & 0 \leqslant s \leqslant p; \\ 0, & \text{其它。} \end{cases} \tag{2.78}$$

则(2.76)可写为

$$\sum_{k=0}^{\infty} \tilde{\theta}_k z^k = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} c_k \tilde{\varphi}_j z^{k+j}. \tag{2.79}$$

令 $l=k+j$，并交换求和次序，(2.79)可写为

$$\sum_{k=0}^{\infty}\tilde{\theta}_k z^k=\sum_{l=0}^{\infty}\left(\sum_{k=0}^{l}c_k\varphi_{l-k}\right)z^l. \tag{2.80}$$

比较(2.80)两边可得

$$\tilde{\theta}_l=\sum_{k=0}^{l}c_k\varphi_{l-k},\qquad l=0,1,2,\cdots.$$

利用 $\varphi_0=1$ 可得

$$c_l=\begin{cases}\theta_l-\sum_{k=0}^{l-1}c_k\varphi_{l-k}, & 0\leqslant l\leqslant q;\\ -\sum_{k=0}^{l-1}c_k\varphi_{l-k}, & q+1\leqslant l.\end{cases} \tag{2.81}$$

由 $c_0=\theta_0$ 开始即可由 ARMA 方程的系数$\{\theta_k\}_0^q$和$\{\varphi_k\}_0^p$用(2.81)式逐步递推而得出全部$\{c_k,k\geqslant 0\}$.

例2.5 试求 ARMA 模型

$$\xi_t-\frac{1}{4}\xi_{t-1}=6\varepsilon_t-5\varepsilon_{t-1}+\varepsilon_{t-2}$$

的 Wold 系数 c_0,c_1,c_2,c_3.

解

$$\varphi_0=1,\qquad \varphi_1=-\frac{1}{4},\qquad \theta_0=6,\qquad \theta_1=-5,\qquad \theta_2=1.$$

由(2.81)递推公式可得

$$c_0=\theta_0=6,$$

$$c_1=\tilde{\theta}_1-c_0\tilde{\varphi}_1=\theta_1-c_0\varphi_1=-5-6\times\left(-\frac{1}{4}\right)=-\frac{7}{2},$$

$$c_2=\theta_2-c_0\tilde{\varphi}_2-c_1\tilde{\varphi}_1=\theta_2-c_1\varphi_1=1-\left(-\frac{7}{2}\right)\times\left(-\frac{1}{4}\right)=\frac{1}{8},$$

$$c_3=-c_2\varphi_1=-\frac{1}{8}\times\left(-\frac{1}{4}\right)=\frac{1}{32}.$$

例2.6 求 ARMA 模型

$$\xi_t - \frac{1}{3}\xi_{t-1} = \frac{2}{3}\varepsilon_t - \frac{1}{3}\varepsilon_{t-1}$$

的 Wold 系数$\{c_k, k\geqslant 0\}$和 Wold 展式。

解

$$\varphi_0 = 1,\quad \varphi_1 = -\frac{1}{3};\quad \theta_0 = \frac{2}{3},\quad \theta_1 = -\frac{1}{3}.$$

$$c_0 = \theta_0 = \frac{2}{3},$$

$$c_1 = \theta_1 - c_0\varphi_1 = -\frac{1}{3} - \frac{2}{3}\times\left(-\frac{1}{3}\right) = -\frac{1}{3^2},$$

$$c_2 = -c_1\varphi_1 = -\left(-\frac{1}{3^2}\right)\times\left(-\frac{1}{3}\right) = -\frac{1}{3^3}.$$

不难看出对一般的 k 有

$$c_k = -c_{k-1}\varphi_1 = -\frac{1}{3^{k+1}}\quad (k\geqslant 1),$$

故 ξ_t 的 Wold 展式为

$$\xi_t = \frac{2}{3}\varepsilon_t - \sum_{k=1}^{\infty}\frac{1}{3^{k+1}}\varepsilon_{t-k}.$$

利用定理 2.3 及以上方法可得以下结果：

定理2.7 设 ξ_t 是 ARMA 序列，其 Wold 系数为$\{c_k\}$，则新息序列$\{\varepsilon_t\}$可表为

$$\varepsilon_t = \sum_{k=0}^{\infty} d_k \xi_{t-k}, \tag{2.82}$$

其中

$$\begin{cases} d_0 = 1/c_0, \\ d_\mu = -d_0\sum_{k=1}^{\mu} c_k d_{\mu-k}, \quad \mu\geqslant 1. \end{cases} \tag{2.83}$$

3° Wold 系数$\{c_k, k\geqslant 0\}$的衰减速度

在 AR(1)模型(2.17)的讨论中，我们指出过 c_k 系数列当$k\to\infty$时是按负指数律衰减的(见(2.21)式)。这一规律在 ARMA 场合下也是正确的。

定理2.8 设 ξ_t 为 ARMA 序列，$\{c_k, k\geqslant 0\}$是它的 Wold 系数，则存在 $\gamma_1, \gamma_2>0$，使

$$|c_k|\leqslant \gamma_1 e^{-k\gamma_2}, \quad k=0,1,2,\cdots. \tag{2.84}$$

并且 ξ_t 的相关函数 $R_\xi(k)$亦呈负指数衰减。

证明 由定理2.2知 ξ_t 的极大解析函数为

$$\Gamma_\xi(z)=\frac{\Theta(z)}{\Phi(z)}=\sum_{k=0}^{\infty} c_k z^k,$$

它在$|z|\leqslant\rho$，$\rho>1$内一致收敛。

$$c_k=\frac{1}{k!}\Gamma_\xi^{(k)}(0), \quad k\geqslant 0. \tag{2.85}$$

由 Cauchy 公式[19]，

$$\Gamma_\xi^{(k)}(z)=\frac{k!}{2\pi i}\int_{\gamma_\rho}\frac{\Gamma_\xi(\zeta)}{(\zeta-z)^{k+1}}\,d\zeta \tag{2.86}$$

其中 γ_ρ 是以原点为圆心以 ρ 为半径的圆周$|\zeta|=\rho$。

令 $\zeta=\rho e^{i\theta}$，$0\leqslant\theta\leqslant 2\pi$，则(2.86)可改写为

$$\Gamma_\xi^{(k)}(0)=\frac{k!}{2\pi i}\int_0^{2\pi}\frac{\Gamma_\xi(\rho e^{i\theta})}{(\rho e^{i\theta})^{k+1}}(i\rho e^{i\theta}d\theta).$$

因而

$$|\Gamma_\xi^{(k)}(0)|\leqslant\frac{k!}{2\pi\rho^k}\int_0^{2\pi}|\Gamma_\xi(\rho e^{i\theta})|d\theta.$$

利用$|\Gamma_\xi(z)|\leqslant M$，则

$$|\Gamma_\xi^{(k)}(0)|\leqslant k!\rho^{-k}M. \tag{2.87}$$

因为$\frac{1}{\rho}<1$，故可选 $\gamma=\ln\rho>0$，使

$$|\Gamma_\xi^{(k)}(0)|\leqslant k!Me^{-\gamma k}.$$

代回(2.85)

$$|c_k|\leqslant Me^{-k\gamma},\qquad k\geqslant 0. \tag{2.88}$$

又

$$\xi_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k},$$

所以

$$E\xi_{t+n}\xi_t=\sum_{k,l=0}^{\infty}c_kc_l\delta_{t+n-k,t-l},\qquad n\geqslant 0.$$

只当 $k=n+l$ 求和项不为零，故

$$R(n)=\sum_{l=0}^{\infty}c_lc_{n+l},\qquad n\geqslant 0. \tag{2.89}$$

从而由(2.88)可得

$$|R(n)|\leqslant\sum_{l=0}^{\infty}|c_l||c_{l+n}|\leqslant\sum_{l=0}^{\infty}Me^{-\gamma l}Me^{-\gamma(l+n)}$$

$$=M^2e^{-n\gamma}/(1-e^{-2\gamma})=\beta e^{-n\gamma},\quad n\geqslant 0, \tag{2.90}$$

其中

$$\beta=\frac{M^2}{1-e^{-2\gamma}}>0.$$

可见 $R(n)$ 随 n 之增长而呈负指数衰减。 ▎

§3 ARMA模型的相关函数与偏相关系数

1° ARMA 模型的相关函数与 Yule-Walker 方程

这一节里将进一步探讨 ARMA 序列的相关函数和谱密度的性质。

定理2.9 设 ξ_t 是 ARMA(p,q)方程

$$\Phi(U)\xi_t=\Theta(U)\varepsilon_t$$

的平稳解，则其相关函数 $R(n)$ 满足以下的差分方程

$$\begin{cases} \sum\limits_{k=0}^{p} \varphi_k R(n-k) = \sum\limits_{k=n}^{q} \theta_k c_{k-n}, & 0 \leqslant n \leqslant q; \\ \sum\limits_{k=0}^{p} \varphi_k R(n-k) = 0, & n > q. \end{cases} \tag{2.91}$$

其中$\{c_k\}$为ξ_t的 Wold 系数。

证明 ARMA 方程

$$\sum_{k=0}^{p} \varphi_k \xi_{n+l-k} = \sum_{k=0}^{q} \theta_k \varepsilon_{n+l-k}$$

两边乘以ξ_l并求期望，得

$$E\left\{ \sum_{k=0}^{p} \varphi_k (\xi_{n+l-k}\xi_l) \right\} = E\left\{ \sum_{k=0}^{q} \theta_k \varepsilon_{n+l-k}\xi_l \right\},$$

即

$$\sum_{k=0}^{p} \varphi_k R(n-k) = \sum_{k=0}^{q} \theta_k (\varepsilon_{n+l-k}, \xi_l)_H. \tag{2.92}$$

然而由$\varepsilon_t \perp H_\xi(t-s), s>0$以及(2.67)式知，当$0 \leqslant k-n$时有

$$(\varepsilon_{n+l-k}, \xi_l) = c_{k-n} \quad (n \leqslant k),$$

其它项为零，故(2.92)可写为

$$\sum_{k=0}^{p} \varphi_k R(n-k) = \sum_{k=n}^{q} \theta_k c_{k-n}. \quad \blacksquare$$

推论 1 设ξ_t为 MA(q)序列，则其相关函数

$$R(n) = \begin{cases} 0, & n > q; \\ \sum\limits_{k=n}^{q} \theta_k \theta_{k-n}, & 0 \leqslant n \leqslant q. \end{cases} \tag{2.93}$$

证明 视 MA(q)序列为 ARMA($0,q$) 序列，则由定理的(2.91)式有($\varphi_0 = 1$)

$$\begin{cases} R(n) = \sum\limits_{k=n}^{q} \theta_k c_{k-n}, & 0 \leqslant n \leqslant q; \\ R(n) = 0, & n > q. \end{cases} \tag{2.94}$$

但是对 MA(q)序列，由等式

$$\xi_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k}=\sum_{k=0}^{q}\theta_k\varepsilon_{t-k}$$

及 H_ξ 中以$\{\varepsilon_t\}$正交基展开的唯一性，

$$\begin{cases}c_k=\theta_k, & k=0,1,2,\cdots,q;\\ c_k=0, & k>q.\end{cases}$$

代回(2.94)即得(2.93)。 ▎

MA(q)序列的相关函数满足(2.93)表明 $R(n)$的截尾性，即

$$R(n)=0,\qquad |n|>q. \tag{2.95}$$

这一性质实质上成为MA序列的特征性刻划，即不仅是MA序列的必要条件，在一定程度上也是充分条件（见下一定理）。

推论 2 设 ξ_t 为AR(p)型序列，则其相关函数 $R(n)$满足以下 Yule-Walker 方程

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p)\\ R(1) & R(0) & \cdots & R(p-1)\\ \cdots & \cdots & \cdots & \cdots\\ R(p) & R(p-1) & \cdots & R(0)\end{bmatrix}\begin{bmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_p\end{bmatrix}=\begin{bmatrix}\theta_0^2\\ \\ 0\\ \\ \end{bmatrix}, \tag{2.96}$$

且

$$R(n)=-\sum_{l=1}^{p}\varphi_l R(n-l),\qquad n\geqslant 1. \tag{2.97}$$

证明 视AR(p)为ARMA(p,0)，则由定理2.8得

$$\begin{cases}\sum\limits_{k=0}^{p}\varphi_k R(n-k)=\theta_0^2, & n=0;\\ \sum\limits_{k=0}^{p}\varphi_k R(n-k)=0, & n>0.\end{cases} \tag{2.98}$$

将(2.98)式写成矩阵形式即为(2.96)。(2.98)式的第二个方程移项后即可得(2.97)式。 ▎

在以后的讨论中，有时也将(2.96)改写为

$$\theta_0^2=\sum_{k=0}^{p}R(k)\varphi_k. \tag{2.99}$$

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \cdots & \cdots & \cdots & \cdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} -\varphi_1 \\ -\varphi_2 \\ \vdots \\ -\varphi_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}. \tag{2.100}$$

以下我们来证明一条非常重要的定理。

定理2.10 设 ξ_t 是平稳列，谱密度 $f_\xi(\lambda)\neq 0, \lambda\in\Pi$，则以下命题等价：

A. ξ_t 是 MA 型序列；

B. ξ_t的相关函数 $R(n)$是截尾型的；

C. ξ_t的谱密度 $f_\xi(\lambda)$可表为

$$f_\xi(\lambda)=\frac{1}{2\pi}|\Theta(e^{-i\lambda})|^2, \qquad \lambda\in\Pi,$$

而 $\Theta(z)$是实系数的多项式，$\Theta(z)\neq 0, |z|\leqslant 1$，$\Theta(0)>0$。

证明

$A\Rightarrow B$ 可见于定理2.8的推论1.

$B\Rightarrow C$. 设 $|R(n)|=0$，$|n|>q$（q 步截尾），由 Wiener--Хинчин 定理

$$f_\xi(\lambda)=\frac{1}{2\pi}\sum_{k=-q}^{q}R(k)e^{-ik\lambda}. \tag{2.101}$$

考虑

$$\hat{f}_\xi(z)=\frac{1}{2\pi}\sum_{k=-q}^{q}R(k)z^k. \tag{2.102}$$

$\hat{f}_\xi(e^{-i\lambda})\neq 0$，并由 $R(-k)=R(k)$知

$$\begin{aligned} \hat{f}_\xi\left(\frac{1}{z_0}\right)&=\frac{1}{2\pi}\sum_{l=-q}^{q}R(-l)z_0^l \\ &=\frac{1}{2\pi}\sum_{l=-q}^{q}R(l)z_0^l=\hat{f}_\xi(z_0), \end{aligned} \tag{2.103}$$

$$\begin{aligned} \hat{f}_\xi\left(\frac{1}{\bar{z}_0}\right)&=\frac{1}{2\pi}\sum_{l=-q}^{q}R(-l)\bar{z}_0^l \\ &=\overline{\left(\frac{1}{2\pi}\sum_{l=-q}^{q}R(l)z_0^l\right)}=\hat{f}_\xi(\bar{z}_0). \end{aligned} \tag{2.104}$$

即有

$$\overline{\hat f_\xi\left(\frac{1}{z_0}\right)} = \overline{\hat f_\xi(z_0)} = \hat f_\xi\left(\frac{1}{\bar z_0}\right) = \hat f_\xi(\bar z_0). \qquad (2.105)$$

可见：若 z_0 为根，则 $z_0,\frac{1}{z_0},\frac{1}{\bar z_0}$ 皆为根（若 z_0 为实根则只有 z_0，$\frac{1}{z_0}$ 两种）。

由于 $\hat f_\xi(z)$ 是 $2q$ 阶多项式，故有 $2q$ 个根。$z_0,(\bar z_0)^{-1},\bar z_0,(z_0)^{-1}$ 可写成

$$z_0,\ (\bar z_0)^{-1};\ (\bar z_0),\ (\bar{\bar z}_0)^{-1}, \qquad (2.106)$$

故 $2q$ 个根可排成（若重根每根算一次全排出来）

$$z_1,\ (\bar z_1)^{-1};\ z_2,\ (\bar z_2)^{-1};\cdots;\ z_q,\ (\bar z_q)^{-1}, \qquad (2.107)$$

其中可要求 $|z_i|>1(i=1,2,\cdots,q)$，因为若 $|z_0|<1$，则 $|\bar z_0^{-1}| = \frac{1}{|z_0|}>1$，二者可互换。

(2.102)式的 $\hat f_\xi(z)$ 可改写为

$$\begin{aligned}\hat f_\xi(z) &= \frac{1}{2\pi z^q}\sum_{k=-q}^{q} R(k) z^{k+q} \\ &= \frac{d}{2\pi z^q}\prod_{k=1}^{q}(z-z_k)\prod_{k=1}^{q}\left(z-\frac{1}{\bar z_k}\right) \\ &= \frac{d}{2\pi}\prod_{k=1}^{q}(z-z_k)\prod_{k=1}^{q}\left(1-\frac{z^{-1}}{\bar z_k}\right),\end{aligned}$$

$$\begin{aligned}\hat f_\xi(e^{-i\lambda}) &= \frac{(-1)^q d}{2\pi\prod_{j=1}^{q}(\bar z_j)}\prod_{k=1}^{q}(e^{-i\lambda}-z_k)\overline{\prod_{k=1}^{q}(e^{-i\lambda}-z_k)} \\ &= \frac{(-1)^q d}{2\pi\prod_{j=1}^{q}(\bar z_j)}\left|\prod_{k=1}^{q}(e^{-i\lambda}-z_k)\right|^2, \qquad \lambda\in\Pi.\end{aligned} \qquad (2.108)$$

若 $\hat f_\xi(e^{-i\lambda})>0$，则

$$D=\frac{(-1)^q d}{\prod_{j=1}^{q}(z_j)}>0。\qquad (2.109)$$

于是可依以下情况定义 $\Theta(z)$：

(1) 若$\{z_k\}$无实数，或虽有实数而

$$\prod_{k=1}^{q}(-z_k)>0,$$

则令

$$\Theta(z)=\sqrt{D}\prod_{k=1}^{q}(z-z_k)。\qquad (2.110)$$

(2) 若$\{z_k\}$有实数而

$$\prod_{k=1}^{q}(-z_k)<0\quad (因|z_k|>1),\qquad (2.111)$$

不妨设 z_q 为实数，令

$$\Theta(z)=\sqrt{D}(z_q-z)\prod_{k=1}^{q-1}(z-z_k)。\qquad (2.112)$$

$$\hat{f}_\xi(e^{-i\lambda})=f_\xi(\lambda)=\frac{1}{2\pi}|\Theta(e^{-i\lambda})|^2,\quad \lambda\in\Pi,$$

并可得以下性质：

a. $\Theta(z)$是 q 阶实系数多项式。

因为 $z_1,z_2,\cdots,z_q$或为实数或其中有成对的共轭复数(见(2.106))，因而可写

$$\prod_{k=1}^{q}(z-z_k)=\prod_{j}(z-z_{k_j})\prod_{l}(z-z_{k_l})(z-\bar{z}_{k_l}),$$

$$(2.113)$$

其中 z_{k_j}为实数，z_{k_l}为复数。然而

$$(z-z_{k_l})(z-\bar{z}_{k_l})=z^2-2\mathrm{Re}(z_{k_l})z+|z_{k_l}|^2$$

是实系数二阶多项式，故(2.113)，进而 $\Theta(z)$皆为实系数多项式。

b. $\Theta(0)>0$。

为此，可分以下几种情况：

(i) 若$\{z_k\}$无实数，则必成对共轭出现，因此

$$\Theta(0)=\sqrt{D}\prod_{l}(-z_{k_l})(-\bar{z}_{k_l})>0。$$

(ii) 若$\{z_k\}$有实数，$\prod_{k=1}^{q}(-z_k)<0$，则由(2.112)

$$\Theta(0)=\sqrt{D}\,z_q\prod_{k=1}^{q-1}(-z_k)>0。$$

(iii) $\{z_k\}$有实数，$\prod_{k=1}^{q}(-z_k)>0$，显然 $\Theta(0)>0$。

至此，B ⇒ C 证完。

C ⇒ A。可由定理 2.4 的推论 3 得出。 |

事实上，我们可以证明以下更一般的定理：

定理2.11 设ξ_t是实平稳列，有谱密度$f_\xi(\lambda)$，则下列命题等价：

A. ξ_t是广 MA 型序列①；

B. ξ_t的相关函数 $R(n)$是截尾型；

C. ξ_t的谱密度$f_\xi(\lambda)$可表为

$$f_\xi(\lambda)=\frac{1}{2\pi}|\Theta(e^{-i\lambda})|^2,\qquad \lambda\in\Pi$$

而 $\Theta(z)$ 是实系数多项式，其根皆在$|z|\geqslant1$（单位圆上及外，且 $\Theta(0)>0$）。

此定理的证明可参照定理2.9，但由于单位圆周上可能有复根，(2.108) 应改为

$$\tilde{f}_\xi(e^{-i\lambda})=\frac{(-1)^n d}{2\pi\prod_{j=1}^{n}z_j}\left|\prod_{k=1}^{n}(e^{-i\lambda}-z_k)\right|^2$$

---

① 指 $\xi_t=\Theta(U)\varepsilon_t$，而 $\Theta(z)$ 的根模$\geqslant1$。

$$\times\prod_{l=1}^{s}(e^{-i\lambda}-e^{-i\alpha_l}-e^{i\alpha_l}+e^{i\lambda}),\qquad(2.114)$$

其中$\{e^{\pm i\alpha_l};\ l=1,2,\cdots,s\}$为单位圆上的根。

利用$\hat f_\xi(e^{-i\lambda})\geqslant0$不难证明(2.114)中每个$\alpha_l$必是偶数个，从而(2.114)右端括号部分可写为

$$\left[\frac{1}{4}|e^{-i\lambda}-e^{i\alpha_l}|^2|e^{-i\lambda}-e^{-i\alpha_l}|^2\right]^{m_l}$$

形式的连乘，因此亦可合并到前部分写成模平方的形式。

关于 ARMA 模型的相关函数有以下特征性的刻划：

定理2.12 设ξ_t是平稳列，谱密度$f_\xi(\lambda)>0$，则它是 ARMA(p,q)型的充分必要条件是：对一切$n>q$的正整数，ξ_t的相关函数$R(n)$满足

$$\sum_{j=0}^{p}\varphi_jR(n-j)=0\ \ (n>q),\qquad(2.115)$$

其中$\Phi(z)=\sum\limits_{j=0}^{p}\varphi_jz^j$是根皆在单位圆外的实系数多项式①。

证明 必须性可见于定理 2.8 中(2.91)式。只须证充分性。令

$$\eta_t=\sum_{j=0}^{p}\varphi_j\xi_{t-j},$$

则

$$\begin{aligned}E(\eta_{t+\tau}\eta_t)&=E\left(\sum_{j=0}^{p}\varphi_j\xi_{t+\tau-j}\right)\left(\sum_{l=0}^{p}\varphi_l\xi_{t-l}\right)\\&=\sum_{l=0}^{p}\varphi_lE\xi_{t-l}\left(\sum_{j=0}^{p}\varphi_j\xi_{t+\tau-j}\right)\\&=\sum_{l=0}^{p}\varphi_l\left(\sum_{j=0}^{p}\varphi_jR(\tau-j+l)\right),\quad\tau\geqslant0.\end{aligned}\qquad(2.116)$$

① 此处要求p,q是最小的阶。

当$\tau>q$时，$\tau+l>q$，由充分性条件，(2.116)括号部分为零，故η_t的相关函数

$$R_{\eta\eta}(\tau)=0.$$

表明η_t是相关函数q步截尾型序列。由定理2.11，它是广MA(q)型，即η_t可表为：

$$\eta_t=\Theta(U)\varepsilon_t \qquad (2.117)$$

但已知$f_\xi(\lambda)$可表为

$$f_\xi(\lambda)=f_\eta(\lambda)/|\Phi(e^{-i\lambda})|^2>0,$$

故$\Theta(z)$在单位圆上无根，表明η_t是MA(q)型。由定理2.4知ξ_t为ARMA(p,q)序列。 ▎

2° ARMA模型的偏相关系数

上一节中我们已证明了刻划MA序列的特征性条件，即MA型的充要条件是相关函数是截尾型。我们继而要问：刻划AR序列的特征性条件是什么？以下用偏相关系数来回答这一问题。

定义2.4 设ξ_t是平稳列，$\{\varphi_j^{(k)};j=1,2,\cdots,k\}$使均方误差

$$\delta_k=\inf_{\alpha_j^{(k)}} E\left|\xi_t-\sum_{j=1}^k\alpha_j^{(k)}\xi_{t-j}\right|^2 =E\left|\xi_t-\sum_{j=1}^k\varphi_j^{(k)}\xi_{t-j}\right|^2 \qquad (2.118)$$

达极小，则称$\varphi_k^{(k)}$为第k步偏相关值，$\{\varphi_k^{(k)};k=1,2,\cdots\}$称为$\xi_t$的偏相关系数。

注1 显见，若令

$$H_{k,t}=L\{\xi_s,\ t-k\leqslant s\leqslant t-1\}^{①}, \qquad (2.119)$$

则$\operatorname{Proj}_{H_{k,t}}\{\xi_t\}$的系数就是$\{\varphi_j^{(k)};j=1,2,\cdots,k\}$。

---

① 指$\{\xi_{t-k},\xi_{t-k+1},\cdots,\xi_{t-1}\}$的线性集。

注 2 求(2.118)的极小值亦可用$\frac{\partial\delta_k}{\partial\alpha_j^{(k)}}=0$的简单方法来求，这时就可导出如下的线性方程组

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(k-1) \\ R(1) & R(0) & \cdots & R(k-2) \\ \cdots & \cdots & \cdots & \cdots \\ R(k-1) & R(k-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} \varphi_1^{(k)} \\ \varphi_2^{(k)} \\ \vdots \\ \varphi_k^{(k)} \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(k) \end{bmatrix}, \tag{2.120}$$

即为 Yule-Walker 方程(2.100)，此中 $\varphi_j^{(k)}=-\varphi_j$。

注 3 (2.120)左端的矩阵

$$\boldsymbol{R}_{k+1}=\begin{bmatrix} R(0) & R(1) & \cdots & R(k) \\ R(1) & R(0) & \ddots & \\ \ddots & & \ddots & \ddots \\ R(k) & \cdots & \cdots & R(0) \end{bmatrix}$$

为 Toeplitz 阵。这种(2.120)线性方程组有递推解法。事实上我们有：

定理2.13 设 ξ_t 为非奇异平稳列，$\boldsymbol{R}_k$ 为其相关阵，则由

$$\boldsymbol{R}_k\begin{bmatrix} \varphi_1^{(k)} \\ \vdots \\ \varphi_k^{(k)} \end{bmatrix}=\begin{bmatrix} R(1) \\ \vdots \\ R(k) \end{bmatrix}$$

确定的$\{\varphi_j^{(k)}; j=1,2,\cdots,k\}$具有以下递推形式

$$\begin{cases} \varphi_1^{(1)}=R(1)/R(0), \\ \varphi_{k+1}^{(k+1)}=\left(R(k+1)-\sum\limits_{j=1}^{k}R(k+1-j)\varphi_j^{(k)}\right) \\ \qquad\qquad \times\left(R(0)-\sum\limits_{j=1}^{k}R(j)\varphi_j^{(k)}\right)^{-1}, \\ \varphi_j^{(k+1)}=\varphi_j^{(k)}-\varphi_{k+1}^{(k+1)}\varphi_{k-(j-1)}^{(k)}, \qquad 1\leqslant j\leqslant k. \end{cases} \tag{2.121}$$

证明 令

$$
\begin{aligned}
a'_k &= (R(1),\cdots,R(k)),\\
\bar{a}'_k &= (R(k),\cdots,R(1)),\\
b'_k &= (\varphi_1^{(k)},\cdots,\varphi_k^{(k)}),\\
c'_k &= (\varphi_1^{(k)},\cdots,\varphi_{k-1}^{(k)}),
\end{aligned}
\qquad (2.122)
$$

$$
T=\begin{bmatrix} 0 & & 1\\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0\end{bmatrix}.
$$

首先指出 $\boldsymbol{R}_k^{-1}$ 是存在的，否则表明必存在常数向量 α 使得

$$
\alpha \boldsymbol{R}_k \alpha' = E\left|\sum_{j=1}^{k}\alpha_j\xi_{k-j+1}\right|^2 = 0,
$$

表明 ξ_l 可用 $\{\xi_1,\cdots,\xi_{l-1}\}$ 准确预报（其中 α_l 是 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 中第一个不为零的系数），则 ξ_l 是奇异的。同理

$$
R(0)-\sum_{j=1}^{k}\varphi_j^{(k)}R(j)\neq 0. \qquad (2.123)
$$

其次不难看出有以下等式成立

$$
\begin{cases}
T^{-1}=T,\\
T\boldsymbol{R}_kT=\boldsymbol{R}_k,\\
Ta_k=\bar{a}_k.
\end{cases} \qquad (2.124)
$$

于是 Yule-Walker 方程可改写为

$$
\boldsymbol{R}_{k+1}b_{k+1}=a_{k+1}
$$

或

$$
\begin{bmatrix}\boldsymbol{R}_k, & \bar{a}_k\\ \bar{a}'_k, & R(0)\end{bmatrix}\begin{bmatrix}c_{k+1}\\ \varphi_{k+1}^{(k+1)}\end{bmatrix}=\begin{bmatrix}a_k\\ R(k+1)\end{bmatrix}.
$$

分块运算后可得

$$
\boldsymbol{R}_kc_{k+1}+\bar{a}_k\varphi_{k+1}^{(k+1)}=a_k, \qquad (2.125)
$$

$$
\bar{a}'_kc_{k+1}+\varphi_{k+1}^{(k+1)}R(0)=R(k+1). \qquad (2.126)
$$

由(2.125),

$$\begin{aligned} c_{k+1} &= R_k^{-1}(a_k - \bar{a}_k \varphi_{k+1}^{(k+1)}) \\ &= b_k - \varphi_{k+1}^{(k+1)} R_k^{-1} \bar{a}_k \\ &= b_k - \varphi_{k+1}^{(k+1)} R_k^{-1} T a_k. \end{aligned} \tag{2.127}$$

然而 $R_k^{-1} = T R_k^{-1} T$，故 $T R_k^{-1} = R_k^{-1} T$，代入(2.127)得

$$c_{k+1} = b_k - \varphi_{k+1}^{(k+1)} T b_k. \tag{2.128}$$

代入(2.126)得

$$R(k+1) = \bar{a}_k' b_k - \bar{a}_k' T b_k \varphi_{k+1}^{(k+1)} + \varphi_{k+1}^{(k+1)} R(0).$$

又由 $\bar{a}_k' = a_k' T$，则

$$\begin{aligned} R(k+1) &= \varphi_{k+1}^{(k+1)}(R(0) - a_k' T T b_k) + \bar{a}_k' b_k \\ &= \varphi_{k+1}^{(k+1)}(R(0) - a_k' b_k) + \bar{a}_k' b_k. \end{aligned}$$

可见

$$\varphi_{k+1}^{(k+1)} = (R(k+1) - \bar{a}_k' b_k)(R(0) - a_k' b_k)^{-1}. \tag{2.129}$$

而(2.129)与(2.128)即为定理所求递推式的矩阵写法。 |

显见，当 ξ_t 为 AR 模型时，它是非奇异的，故定理 2.12 的递推公式可用。

为了讲清楚偏相关系数的概念，以下介绍一条非常有用的引理。

引理2.1 设 $y=(y_1,\cdots,y_n)'$ 与 x 皆为实随机变量，Eyy' 非退化，则

$$\operatorname*{Proj}_{\mathscr{H}(y)}(x) = (Exy')(Eyy')^{-1}y. \tag{2.130}$$

证明 问题等价于求使

$$E(x-a'y)(x-a'y)'$$

达最小的系数 a。设

$$\begin{aligned} \delta_a &= E(x-a'y)(x-a'y)' \\ &= E(x^2) - a'E(xy) - E(xy')a + a'E(yy')a. \end{aligned}$$

利用矩阵微商公式(见[20])，Eyy' 为对称阵，则

$$\frac{\partial \delta_a}{\partial a} = -E(xy) - E(xy) + 2(Eyy')a = 0,$$

即

$$\alpha=(Eyy')^{-1}Exy, \tag{2.131}$$

从而

$$\operatorname*{Proj}_{\mathscr{L}(y)}(x)=\alpha' y=(Exy')(Eyy')^{-1}y. \quad \blacksquare$$

若改 x 为向量，则对每个分量求投影，因此(2.130)亦成立。

定理2.14 设 ξ_t 是 ARMA 模型，令

$$\begin{aligned}\tilde{\xi}_t&=\xi_t-\operatorname*{Proj}_{H_{k-1,t}}(\xi_t),\\ \tilde{\xi}_{t-k}&=\xi_{t-k}-\operatorname*{Proj}_{H_{k-1,t}}(\xi_{t-k}),\end{aligned} \tag{2.132}$$

则

$$\varphi_k^{(k)}=\frac{E\tilde{\xi}_t\tilde{\xi}_{t-k}}{\sqrt{E|\tilde{\xi}_t|^2}\sqrt{E|\tilde{\xi}_{t-k}|^2}}, \tag{2.133}$$

其中 $H_{k,t}$ 由(2.119)定义。

证明 令

$$x=(\xi_t,\xi_{t-k})',\qquad y=(\xi_{t-1},\cdots,\xi_{t-k+1})',$$

由引理2.1知

$$\operatorname*{Proj}_{\mathscr{L}(y)}(x)=(Exy')(Eyy')^{-1}y.$$

令

$$\hat{x}=\operatorname*{Proj}_{\mathscr{L}(y)}(x),\qquad \tilde{x}=x-\hat{x}\triangleq(\tilde{\xi}_t,\tilde{\xi}_{t-k})',$$

则

$$\begin{aligned}E\tilde{x}\tilde{x}'&=E(x-\hat{x})x'-E(x-\hat{x})\hat{x}'\\&=Exx'-(Exy')(Eyy')^{-1}(Eyx').\end{aligned} \tag{2.134}$$

又（符号参看(2.122)）

$$Exx'=\begin{pmatrix}R(0) & R(k)\\ R(k) & R(0)\end{pmatrix},$$

$$Exy'=\begin{pmatrix}R(1) & R(2) & \cdots & R(k-1)\\ R(k-1) & R(k-2) & \cdots & R(1)\end{pmatrix}=\begin{pmatrix}\alpha'_{k-1}\\ \bar{\alpha}'_{k-1}\end{pmatrix}.$$

而

$$Eyy' = \begin{pmatrix} R(0) & R(1) & \cdots & R(k-2) \\ \vdots & & & \vdots \\ R(k-2) & \cdots & \cdots & R(0) \end{pmatrix} = \boldsymbol{R}_{k-1},$$

将以上各式代入(2.134)可得（引用(2.122)符号）:

$$E\tilde{x}\tilde{x}' = \begin{pmatrix} R(0) & R(k) \\ R(k) & R(0) \end{pmatrix} - \begin{pmatrix} a'_{k-1} \\ \bar{a}'_{k-1} \end{pmatrix} \boldsymbol{R}_{k-1}^{-1}(a_{k-1} \quad \bar{a}_{k-1})$$

$$= \begin{pmatrix} R(0) & R(k) \\ R(k) & R(0) \end{pmatrix} - \begin{pmatrix} a'_{k-1} \\ \bar{a}'_{k-1} \end{pmatrix} (b_{k-1} \quad Tb_{k-1})$$

$$= \begin{pmatrix} R(0) - a'_{k-1}b_{k-1} & R(k) - \bar{a}'_{k-1}b_{k-1} \\ R(k) - \bar{a}'_{k-1}b_{k-1} & R(0) - a'_{k-1}b_{k-1} \end{pmatrix}. \quad (2.135)$$

另一方面由定义

$$E\tilde{x}\tilde{x}' = E\begin{pmatrix} \tilde{\xi}_t^2 & \tilde{\xi}_t\tilde{\xi}_{t-k} \\ \tilde{\xi}_t\tilde{\xi}_{t-k} & \tilde{\xi}_{t-k}^2 \end{pmatrix}, \quad (2.136)$$

由(2.135)和(2.136)比较，则

$$\frac{E\tilde{\xi}_t\tilde{\xi}_{t-k}}{\sqrt{E|\tilde{\xi}_t|^2} \sqrt{E|\tilde{\xi}_{t-k}|^2}} = \frac{R(k) - \bar{a}'_{k-1}b_{k-1}}{R(0) - a'_{k-1}b_{k-1}}. \quad (2.137)$$

将(2.137)右端与定理2.12的(2.129)式相比知

$$\frac{E\tilde{\xi}_t\tilde{\xi}_{t-k}}{\sqrt{E|\tilde{\xi}_t|^2} \sqrt{E|\tilde{\xi}_{t-k}|^2}} = \varphi_k^{(k)}. \quad \blacksquare$$

定理 2.13 的结果表明 $\varphi_k^{(k)}$ 是随机变量 ξ_t, ξ_{t-k} 各自去掉 $\{\xi_{t-1}, \xi_{t-2}, \cdots, \xi_{t-k+1}\}$ 的影响（表现为投影分量）后，残余分量之间的相关系数，故称为偏相关系数。

以下定理刻划了自回归 AR 模型的特征:

定理2.15 设 ξ_t 是非奇异的平稳列，则它是 AR(p) 模型的充分必要条件是 ξ_t 的偏相关系数 $\varphi_k^{(k)}$ 是 p 步截尾的，即

$$\varphi_k^{(k)}=\begin{cases}0, & k>p;\\ \varphi_p^{(p)}\neq 0, & k=p.\end{cases} \tag{2.138}$$

证明 先证必需性。

若已知ξ_t为AR(p)型，设其模型方程为

$$\Phi(U)\xi_t=\theta_0\varepsilon_t, \qquad \theta_0>0,$$

则由定义知$\varphi_k^{(k)}$是使

$$\delta_k=E\left[\xi_t-\sum_{j=1}^{k}\varphi_j^{(k)}\xi_{t-j}\right]^2 \tag{2.139}$$

极小化而得到的系数$\{\varphi_j^{(k)};\ j=1,2,\cdots,k\}$中的最后一个。对$k\geqslant p$，由模型方程，可将(2.139)改写为

$$\begin{aligned}\delta_k&=E\left[-\sum_{j=1}^{p}\varphi_j\xi_{t-j}+\theta_0\varepsilon_t-\sum_{j=1}^{k}\varphi_j^{(k)}\xi_{t-j}\right]^2\\&=E\left[\theta_0\varepsilon_t-\sum_{j=1}^{p}(\varphi_j+\varphi_j^{(k)})\xi_{t-j}-\sum_{j=p+1}^{k}\varphi_j^{(k)}\xi_{t-j}\right]^2.\end{aligned} \tag{2.140}$$

显见(2.140)的后两项属于$H_\xi(t-1)$，而$\varepsilon_t\perp H_\xi(t-1)$，故

$$\delta_k=\theta_0^2+E\left[\sum_{j=1}^{p}(\varphi_j+\varphi_j^{(k)})\xi_{t-j}+\sum_{j=p+1}^{k}\varphi_j^{(k)}\xi_{t-j}\right]^2\geqslant\theta_0^2. \tag{2.141}$$

而(2.141)右端等号成立当且仅当

$$\varphi_j^{(k)}=\begin{cases}-\varphi_j, & 1\leqslant j\leqslant p;\\ 0, & p<j\leqslant k.\end{cases} \tag{2.142}$$

"当"是显然的，而"仅当"是利用了ξ_t的非奇异性。当$k=p$，由(2.142)知$\varphi_p^{(p)}=-\varphi_p\neq 0$，而$\varphi_k^{(k)}=0$，当$k>p$，即$\{\varphi_k^{(k)}\}$是$p$步截尾的。

反之，若ξ_t的偏相关系数是p步截尾的，则由定理2.12的递推公式（见(2.121)），当$k>p$时：

$$\varphi_j^{(k+1)}=\varphi_j^{(k)}-\varphi_{k+1}^{(k+1)}\varphi_{k-(j-1)}^{(k)}$$
$$=\varphi_j^{(k)}\ (\text{因为}\ \varphi_{k+1}^{(k+1)}=0),\quad j=1,2,\cdots,k.$$

再递推

$$\varphi_j^{(k+1)}=\varphi_j^{(k-l)}=\cdots=\varphi_j^{(p)}\ (0\leqslant l\leqslant k-p). \tag{2.143}$$

于是，对$1\leqslant j\leqslant p$，

$$\varphi_j^{(p)}=\varphi_j^{(k+1)}. \tag{2.144}$$

对 $p<j\leqslant k$，取 $l=k-j$，则

$$\varphi_j^{(k+1)}=\varphi_j^{(k-l)}=\varphi_j^{(j)}=0\ (j>p). \tag{2.145}$$

因此，由递推公式

$$0=\varphi_{k+1}^{(k+1)}=\frac{\left(R(k+1)-\sum_{j=1}^{k}R(k+1-j)\varphi_j^{(k)}\right)}{\left(R(0)-\sum_{j=1}^{k}R(j)\varphi_j^{(k)}\right)}\quad(k>p)$$

可得

$$R(k+1)=\sum_{j=1}^{k}R(k+1-j)\varphi_j^{(k)}. \tag{2.146}$$

又由(2.144)和(2.145)则当 $k\geqslant p$ 时(2.146)可改写为

$$R(k+1)=\sum_{j=1}^{p}R(k+1-j)\varphi_j^{(p)}. \tag{2.147}$$

然而，当 $0\leqslant k\leqslant p$ 时，(2.147)式也是成立的，因为 $\{\varphi_j^{(p)}\}$由定义2.4的注解可知应满足 Y-W 方程

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ \cdots & \cdots & \cdots & \cdots \\ R(k) & R(k-1) & \cdots & R(p-k-1) \\ \cdots & \cdots & \cdots & \cdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix}\begin{bmatrix}\varphi_1^{(p)} \\ \vdots \\ \varphi_{k+1}^{(p)} \\ \vdots \\ \varphi_p^{(p)}\end{bmatrix}$$

$$
=\begin{bmatrix} R(1) \\ \vdots \\ R(k+1) \\ \vdots \\ R(p) \end{bmatrix}.
$$

可见有

$$
R(k)-\sum_{l=1}^{p} R(k-l)\varphi_l^{(p)}=0, \quad k\geqslant 1. \tag{2.148}
$$

令

$$
\Phi(U)=1-\sum_{l=1}^{p}\varphi_l^{(p)}U^l,
$$

则(2.148)可写成

$$
\Phi(U)R(k)=0, \quad k\geqslant 1. \tag{2.149}
$$

令 $\eta_t=\Phi(U)\xi_t$，则对 $\tau>0$有：

$$
\begin{aligned}
E\eta_t\eta_{t-k} &= E[\Phi(U)\xi_t\Phi(U)\xi_{t-k}] \\
&= \sum_{j=0}^{p}\left\{\sum_{i=0}^{p}\varphi_i^{(p)}R(k+j-i)\right\}\varphi_j^{(p)} \\
&= \sum_{j=0}^{p}\varphi_j^{(p)}(\Phi(U)R(k+j))=0 \quad (\text{由}(2.149)).
\end{aligned}
$$

表明 η_t 是不相关列。

记

$$
\begin{aligned}
\theta_0^2=E|\eta_t|^2 &= E|\xi_t-\sum_{l=1}^{p}\varphi_l^{(p)}\xi_{t-l}|^2 \\
&\geqslant C_0^2 \quad (\text{见}(2.75)\text{式}) \\
&>0 \quad (\xi_t\text{是非奇异的}),
\end{aligned}
$$

于是可令 $\varepsilon_t=\eta_t/\theta_0$，则 ε_t 是标准不相关列，并有

$$
\Phi(U)\xi_t=\theta_0\varepsilon_t \tag{2.150}
$$

成立。最后还需证(2.150)中的 $\Phi(z)$的根皆在单位圆外。

反证，设 $\Phi(z)$有 μ^{-1} 的根不在单位圆外，则$|\mu|\geqslant 1$ ($\Phi(z)$

不会有零根)。

记 $\Phi(z)=(1-\mu z)\phi_1(z)$，其中 $\phi_1(z)$ 为 $p-1$ 次的多项式，令

$$Z(t)=\phi_1(U)\xi_t,$$

则

$$\theta_0\varepsilon_t=\Phi(U)\xi_t=Z(t)-\mu Z(t-1). \tag{2.151}$$

$$\begin{aligned}E\{\varepsilon_t\xi_{t-k}\}&=\frac{1}{\theta_0}E(\Phi(U)\xi_t\xi_{t-k})\\&=\frac{1}{\theta_0}\sum_{l=0}^{p}\varphi_l^{(p)}E(\xi_{t-l}\xi_{t-k})\\&=\frac{1}{\theta_0}\Phi(U)R(k)=0,\qquad k\geqslant 1,\end{aligned}$$

可见

$$\varepsilon_t\perp\xi_{t-k},\qquad k\geqslant 1.$$

从而

$$Z(t-1)=\phi_1(U)\xi_{t-1}\perp\varepsilon_t.$$

则(2.151)为

$$\begin{aligned}E[Z(t)]^2&=E[\theta_0\varepsilon_t\oplus\mu Z(t-1)]^2\\&=\theta_0^2+\mu^2E[Z(t-1)]^2\\&>\mu^2E[Z(t-1)]^2\\&\geqslant E[Z(t-1)]^2.\end{aligned} \tag{2.152}$$

由 ξ_t 平稳，$Z(t)$ 亦平稳，故(2.152)矛盾。 ▌

§4　时间序列的马氏扩张问题

1° ARMA 模型与 Markov 性质

设 y_t 是一维平稳列，利用§2中符号，记

$$H_y(t)=\mathscr{L}\{y_\tau|\tau\leqslant t\}.$$

显见

$$\mathscr{L}\{y_t\}\subset H_y(t)\subset H_y(t+\tau),\quad \tau>0.$$

如果有等式

$$\operatorname*{Proj}_{H_y(t)}(y_t+1)=\operatorname*{Proj}_{\mathscr{L}\{y_t\}}(y_t+1) \tag{2.153}$$

成立，则称 y_t 是平稳马氏序列。

设 $\boldsymbol{y}_t$ 是 p 维过程，令

$$H_y=\mathscr{L}\{\boldsymbol{y}_i(t);\ i=1,2,\cdots,p,t=0,\pm1,\cdots\} \tag{2.154}$$

$$H_y(t)=\mathscr{L}\{\boldsymbol{y}_i(\tau);\ \tau\leqslant t,\ i=1,2,\cdots,p\}, \tag{2.155}$$

$$\mathscr{L}\{\boldsymbol{y}_t\}=\mathscr{L}\{y_i(t);\ i=1,2,\cdots,p\}. \tag{2.156}$$

若 $\boldsymbol{y}_t$ 是平稳列，具有

$$\operatorname*{Proj}_{H_y(t)}(\boldsymbol{y}_{t+1})=\operatorname*{Proj}_{\mathscr{L}(\boldsymbol{y}_t)}(\boldsymbol{y}_{t+1})\ (\text{分别投影}), \tag{2.157}$$

则称 $\boldsymbol{y}_t$ 具有马氏性（或 p 维平稳马氏序列）。

从(2.157)式马氏性的定义中知道这类序列具有很深刻的意义，即用$\{\boldsymbol{y}(\tau)|\tau\leqslant t\}$的观察来了解未来 $\boldsymbol{y}(t+1)$，只需要用到“现在”—— t 时刻的资料$\{\boldsymbol{y}(t)\}$而与“过去”的资料$\{\boldsymbol{y}(\tau)|\tau<t\}$无关。然而以下将证明，在一定程度上一维的ARMA模型必是这种多维马氏序列的分量，由此揭示了ARMA模型（或具有理谱密度）序列与马氏性的内在联系，这一工作属于Doob（见[13]）。

定理2.16 设 $x(t)$是平稳列，其谱密度$f_x(\lambda)>0$，则它是ARMA(p,q)，$q\leqslant p-1$，模型的充分必要条件是：$x(t)$是下述 p 维马氏自回归平稳列

$$\boldsymbol{Y}(t)=\boldsymbol{F}\boldsymbol{Y}(t-1)+\boldsymbol{\eta}(t) \tag{2.158}$$

的一个分量，矩阵 $\boldsymbol{F}$ 的特征根的模皆小于1，而 $\boldsymbol{\eta}(t)$ 是 p 维不相关列：

$$E\boldsymbol{\eta}(t)\equiv0,$$

$$E\boldsymbol{\eta}(t)\boldsymbol{\eta}'(s)=\boldsymbol{G}\delta_{t,s}, \tag{2.159}$$

其中 $\boldsymbol{G}$ 为 $p\times p$ 的方阵，并且

$$E\boldsymbol{\eta}(t)\boldsymbol{Y}(t-l)\equiv 0,\quad l\geqslant 1. \tag{2.160}$$

证明 首先证充分性。

设 $\boldsymbol{F}$ 的特征多项式为

$$\det[\boldsymbol{F}-\lambda \boldsymbol{I}]=\sum_{j=0}^{p}\varphi_j\lambda^{p-j}.$$

由 $\boldsymbol{Y}(t)=\boldsymbol{F}\boldsymbol{Y}(t-1)+\boldsymbol{\eta}(t)$，则

$$\begin{aligned}\boldsymbol{Y}(t-j)&=\boldsymbol{F}^2\boldsymbol{Y}(t-j-2)+\boldsymbol{F}\boldsymbol{\eta}(t-j-1)+\boldsymbol{\eta}(t-j)\\&=\boldsymbol{F}^{p-j}\boldsymbol{Y}(t-p)+\sum_{i=j}^{p-1}\boldsymbol{F}^{i-j}\boldsymbol{\eta}(t-i).\end{aligned}$$

所以

$$\begin{aligned}\sum_{j=0}^{p}\varphi_j\boldsymbol{Y}(t-j)&=\sum_{j=0}^{p}\varphi_j\boldsymbol{F}^{p-j}\boldsymbol{Y}(t-p)\\&\quad+\sum_{j=0}^{p-1}\varphi_j\sum_{i=j}^{p-1}\boldsymbol{F}^{i-j}\boldsymbol{\eta}(t-i)\\&=\left(\sum_{j=0}^{p}\varphi_j\boldsymbol{F}^{p-j}\right)\boldsymbol{Y}(t-p)\\&\quad+\sum_{i=0}^{p-1}\sum_{j=0}^{i}\varphi_j\boldsymbol{F}^{i-j}\boldsymbol{\eta}(t-i).\end{aligned} \tag{2.161}$$

利用线性代数中的 Hamilton-Cally 定理(见[21])，$\boldsymbol{F}$应是它的特征多项式

$$\det[\boldsymbol{F}-\lambda \boldsymbol{I}]=\sum_{j=0}^{p}\varphi_j\lambda^{p-j}$$

的根，因而(2.161)右端第一项为零。

$$\begin{aligned}\sum_{j=0}^{p}\varphi_j\boldsymbol{Y}(t-j)&=\sum_{i=0}^{p-1}\left[\sum_{j=0}^{i}\varphi_j\boldsymbol{F}^{i-j}\right]\boldsymbol{\eta}(t-i)\\&\overset{\text{记}}{=}\sum_{i=0}^{p-1}\boldsymbol{B}_i\boldsymbol{\eta}(t-i),\end{aligned} \tag{2.162}$$

其中

$$B_l=\sum_{j=0}^{l}\varphi_j F^{l-j}=(B_{l,\nu}^{(l)}).\qquad(2.163)$$

假定 $x(t)$ 是对应于 $Y(t)$ 的第 l 个分量，将(2.162) 两边对应于第 l 个分量取出，则可得

$$\sum_{j=0}^{p}\varphi_j x(t-j)=\sum_{j=0}^{p-1}B_l^{(j)}\eta(t-j),\qquad(2.164)$$

其中 $B_l^{(j)}=(B_{l,1}^{(j)},B_{l,2}^{(j)},\cdots,B_{l,p}^{(j)})$.

令 $\zeta(t)=\sum_{j=0}^{p-1}B_l^{(j)}\eta(t-j)$ (它是一维的)，则

$$E\zeta(t)=0\quad(\text{因}Ex(t)\equiv 0),$$

$$R(n,m)=E[\zeta(n)\zeta(m)]$$

$$=\sum_{i,j=0}^{p-1}\sum_{\mu,\nu=1}^{p}B_{l,\mu}^{(i)}B_{l,\nu}^{(j)}E[\eta_\mu(n-i)\eta_\nu(m-j)]$$

或

$$R(m+s,m)=\sum_{i,j=0}^{p-1}\sum_{\mu,\nu=1}^{p}B_{l,\mu}^{(i)}B_{l,\nu}^{(j)}E[\eta_\mu(m+s-i)\eta_\nu(m-j)].$$

(2.165)

利用(2.159)上式只当 $m+s-i=m-j$ 有值，由此可得 (2.165) 与 m 无关，即 $\zeta(t)$ 是平稳列，并且

$$R_\zeta(s)=R(m+s,m)$$

$$=\begin{cases}\displaystyle\sum_{\mu,\nu=1}^{p}\sum_{j=0}^{p-1}B_{l,\mu}^{(s+j)}B_{l,\nu}^{(j)}g_{\mu,\nu}, & 0\leqslant|s|\leqslant p-1;\\ 0, & |s|\geqslant p.\end{cases}$$

(2.166)

其中 $(g_{\mu,\nu})=G$. 可见 $\zeta(t)$ 的相关函数是截尾型，由定理 2.10，则 $\zeta(t)$ 必是广 MA 序列，其阶数 $q\leqslant p-1$.

从而(2.164)可表为

$$\sum_{j=0}^{p} \varphi_j x(t-j) = \sum_{s=0}^{p-1} \theta_s \varepsilon(t-s), \quad \theta_0 > 0, \tag{2.167}$$

并且$\sum_{s=0}^{p-1} \theta_s z^s = \Theta(z)$的根的模皆不大于 1 .

以下证明在单位圆上$\Theta(z)$不会有根，事实上不难证明有

$$\frac{1}{2\pi} |\Theta(e^{-i\lambda})|^2 = f_\zeta(\lambda) = \left| \sum_{j=0}^{p} \varphi_j e^{-ij\lambda} \right|^2 f_x(\lambda). \tag{2.168}$$

由条件知$\sum_{j=0}^{p} \varphi_j z^{p-j}$的根模皆小于 1 ，则$\sum_{j=0}^{p} \varphi_j z^j$的根的模皆大于 1 ，又由$f_x(\lambda)>0$，则$|\Theta(e^{-i\lambda})|^2>0$对$-\pi\leqslant\lambda\leqslant\pi$成立，故(2.167)是 ARMA 模型.

以下证必须性.

设$x(t)$满足 ARMA 方程

$$\sum_{j=0}^{p} \varphi_j x(t-j) = \sum_{l=0}^{q} \theta_l \varepsilon(t-l), \quad \theta_0 > 0. \tag{2.169}$$

设$H_x(n) = \mathscr{L}\{x(t)|t\leqslant n\}$. 简记

$$P_n(x(s)) = \underset{H_x(n)}{\mathrm{Proj}}(x(s)). \tag{2.170}$$

令

$$\boldsymbol{Y}(n) = P_n \begin{bmatrix} x(n) \\ x(n+1) \\ \vdots \\ x(n+p-1) \end{bmatrix}. \tag{2.171}$$

则$\boldsymbol{Y}(n)$为p维平稳列，其分量（利用$x(t)$的 Wold 分解）

$$y^{(l+1)}(n) = P_n(x(n+l)) = \sum_{k=l}^{\infty} c_k \varepsilon(n-k+l). \tag{2.172}$$

则

(1) $\{y^{(l+1)}(n);\ n=0,\pm1,\cdots\}$是平稳列。

(2) $H_x(n)=H_Y(n)$。

设$\zeta\in H_x(n)$，则

$$\zeta=\lim_m\sum_{k=1}^m a_k^{(m)}x(k_m)\quad (k_m\leqslant n)$$

可改写为

$$\zeta=\lim_m\sum_{k=1}^m(\boldsymbol{\beta}_k^{(m)})'\boldsymbol{Y}(k_m)\in H_Y(n),$$

其中$k_m\leqslant n$，$\boldsymbol{\beta}_k^{(m)}=(a_k^{(m)},0,\cdots,0)'$。反之，若$\zeta\in H_Y(n)$，由$\boldsymbol{Y}(n)$的每个分量皆属于$H_x(n)$（由投影定义），故$\zeta\in H_x(n)$。

由(2)知

$$\operatorname*{Proj}_{H_x(n)}=\operatorname*{Proj}_{H_Y(n)}=P_n. \tag{2.173}$$

(3) $P_n(P_{n+1})=P_n$。因为$H_x(n+1)\supset H_x(n)$。

记

$$\begin{aligned}\boldsymbol{Y}(n)&=(y_1(n),y_2(n),\cdots,y_p(n))'\\&=P_n(x(n),x(n+1),\cdots,x(n+p-1))'.\end{aligned} \tag{2.174}$$

则

$$\begin{aligned}P_n\boldsymbol{Y}(n+1)&=P_nP_{n+1}(x(n+1),x(n+2),\cdots,x(n+p))'\\&=P_n(x(n+1),\ x(n+2),\cdots,x(n+p))'.\end{aligned} \tag{2.175}$$

将(2.174)和(2.175)比较并用(2.169)知

$$P_nx(n+l)=y_{l+1}(n),\qquad l=1,2,\cdots,p-1. \tag{2.176}$$

$$\begin{aligned}P_nx(n+p)&=P_n\Big(-\sum_{j=1}^p\varphi_jx(n+p-j)+\sum_{l=0}^q\theta_l\varepsilon(n+p-l)\Big)\\&=-\sum_{j=1}^p\varphi_jy_{p-j+1}(n)+\sum_{l=0}^q\theta_lP_n\varepsilon(n+p-l).\end{aligned} \tag{2.177}$$

由于$\varepsilon(n+p-q)\perp H_x(n)\quad(p-q\geqslant 1)$，故(2.177)的第二求和项为零，因此可得(2.175)的改写式：

$$P_nY(n+1)=\left(y_2(n),\cdots,y_p(n),-\sum_{j=1}^{p}\varphi_j y_{p-j+1}(n)\right)'.$$

即

$$P_nY(n+1)=\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ & 0 & 1 & \cdots & 0\\ & & \ddots & \ddots & \\ & 0 & & \ddots & \\ & & & \ddots & 1\\ -\varphi_p & -\varphi_{p-1} & \cdots & -\varphi_2 & -\varphi_1\end{pmatrix}\begin{pmatrix}y_1(n)\\ y_2(n)\\ \vdots\\ y_p(n)\end{pmatrix}\tag{2.178}$$

$$=FY(n).\tag{2.179}$$

其中F为(2.178)右端的$p\times p$方阵。

令

$$\eta(n+1)=Y(n+1)-FY(n).\tag{2.180}$$

以下证明$\{\eta(n)\}$是p维不相关列。

首先可看出

$$\begin{aligned}\eta(n)&=Y(n)-FY(n-1)\\&=Y(n)-P_{n-1}Y(n)\\&=P_n\begin{bmatrix}x(n)\\x(n+1)\\\vdots\\x(n+p-1)\end{bmatrix}-P_{n-1}\begin{bmatrix}x(n)\\x(n+1)\\\vdots\\x(n+p-1)\end{bmatrix}\\&=(P_n-P_{n-1})\begin{bmatrix}x(n)\\x(n+1)\\\vdots\\x(n+p-1)\end{bmatrix}\end{aligned}$$

$$=\begin{bmatrix} c_0\varepsilon(n) \\ c_1\varepsilon(n) \\ \vdots \\ c_{p-1}\varepsilon(n) \end{bmatrix} \text{(利用 Wold 分解性质)}, \tag{2.181}$$

$$E\eta(n)\eta'(m)=(c_i c_k)_{0\le i,k\le p-1}E\varepsilon(n)\varepsilon(m)$$
$$=G\delta_{n,m}, \tag{2.182}$$

其中由 $c_0^2>0$，G 是非零阵，$\{\eta(n)\}$是不相关列。

$$P_n\eta(n+1)=P_nY(n+1)-FP_nY(n)$$
$$=P_nY(n+1)-FY(n)=0, \tag{2.183}$$

表明

$$\eta(n+1)\perp H_Y(n-k),\quad k\geqslant 0. \tag{2.184}$$

最后再证明 F 的特征根的模皆小于 1。事实上

$$\det[F-\lambda I]=\det\begin{bmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ & -\lambda & 1 & \cdots & 0 \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ 0 & & & -\lambda & 1 \\ -\varphi_p & -\varphi_{p-1} & \cdots & -\varphi_2 & -\lambda-\varphi_1 \end{bmatrix}$$

$$=(-\lambda)\begin{vmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ & -\lambda & 1 & \cdots & 0 \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ 0 & & & -\lambda & 1 \\ -\varphi_{p-1} & & \cdots & & -\lambda-\varphi_1 \end{vmatrix}$$

$$+(-1)^{p-1}(-\varphi_p)\begin{vmatrix} 1 & & & 0 \\ -\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots \\ 0 & & -\lambda & 1 \end{vmatrix}$$

$$= (-\lambda)^{p-2}\begin{vmatrix} -\lambda & 1 \\ -\varphi_2 & -\lambda-\varphi_1 \end{vmatrix} + (-1)^p \varphi_3 \lambda^{p-3}$$

$$+ \cdots + (-1)^p \varphi_p$$

$$= (-1)^{p-2} \sum_{j=0}^{p} \varphi_j \lambda^{p-j} \quad (\varphi_0 = 1). \tag{2.185}$$

然而由 ARMA 模型(2.169) $\sum_{j=0}^{p} \varphi_j \lambda^j$ 的根模皆大于 1，则(2.185)右边多项式的根模皆小于1。定理得证。 ▎

2° 马氏扩张的进一步讨论

上述定理从理论上指出了 ARMA 模型与马氏性的内在关系。以下定理在稍为具体的条件下给出了如何进行马氏扩张的构造性方法。

定理2.17 设 $x(t)$是宽平稳零均值的序列，具有有理谱密度，即ARMA$(n, n-1)$：

$$x(t) = \int_{\Pi} e^{i\lambda t} \frac{P_{n-1}(e^{-i\lambda})}{Q_n(e^{-i\lambda})} dz_\varepsilon(\lambda), \tag{2.186}$$

其中 $dz_\varepsilon(\lambda)$是白噪声 $\varepsilon(t)$对应的随机正交测度

$$\varepsilon(t) = \int_{\Pi} e^{i\lambda t} dz_\varepsilon(\lambda), \tag{2.187}$$

而

$$P_{n-1}(z) = \sum_{k=0}^{n-1} \theta_k z^k, \tag{2.188}$$

$$Q_n(z) = \sum_{k=0}^{n} \varphi_k z^k. \tag{2.189}$$

令

$$\begin{cases} c_0 = \theta_0; \\ c_k = \theta_k - \sum_{l=1}^{n} \varphi_l c_{k-l}, \quad 1 \leqslant k \leqslant n-1; \\ c_k = -\sum_{l=1}^{n} \varphi_l c_{k-l}, \quad n \leqslant k. \end{cases} \tag{2.190}$$

$$\begin{cases} w_1(e^{-i\lambda}) = P_{n-1}(e^{-i\lambda})/Q_n(e^{-i\lambda}), \\ w_{k+1}(e^{-i\lambda}) = e^{i\lambda}(w_k(e^{-i\lambda}) - c_{k-1}), \ k = 1,2,\cdots,n-1. \end{cases} \tag{2.191}$$

以及

$$y_k(t) = \int_{\Pi} e^{i\lambda t} w_k(e^{-i\lambda}) dZ_\varepsilon(\lambda), \quad k = 1,2,\cdots,n. \tag{2.192}$$

$$\boldsymbol{Y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix}. \tag{2.193}$$

则

(1) $Ey_j(s)\overline{\varepsilon(t)} = 0, \quad s < t, \ j = 1,2,\cdots,n.$ (2.194)

(2) $x(t) = y_1(t)$，并且 $\boldsymbol{Y}(t)$ 是马氏序列，

$$\boldsymbol{Y}(t+1) = \boldsymbol{A}\boldsymbol{Y}(t) + \boldsymbol{C}\varepsilon(t+1), \tag{2.195}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ & 0 & 1 & \cdots & 0 \\ & 0 & & \ddots & \\ & & & & 1 \\ -\varphi_n & \cdots & \cdots & \cdots & -\varphi_1 \end{bmatrix} \tag{2.196}$$

证明 本定理是定理2.15的具体化。由(2.171)知

$$\boldsymbol{Y}(t) = P_t \begin{bmatrix} x(t) \\ x(t+1) \\ \vdots \\ x(t+n-1) \end{bmatrix} = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}. \tag{2.197}$$

但由 ARMA 模型知

$$P_t x(t+l)=\sum_{k=l}^{\infty} c_k \varepsilon_{t+l-k}, \qquad l=0,1,2,\cdots,n-1. \tag{2.198}$$

故

$$P_t x(t+l)=\int_{\Pi} e^{i\lambda t}\left(\sum_{k=l}^{\infty} c_k e^{-ik\lambda}\right) e^{il\lambda} dZ_{\varepsilon}(\lambda)$$

$$\triangleq \int_{\Pi} e^{i\lambda t} W_{l+1}(e^{-i\lambda}) dZ_{\varepsilon}(\lambda). \tag{2.199}$$

但是

$$W_1(e^{-i\lambda})=\sum_{k=0}^{\infty} c_k e^{-ik\lambda}=\frac{P_{n-1}(e^{-i\lambda})}{Q_n(e^{-i\lambda})},$$

$$W_{l+1}(e^{-i\lambda})=e^{il\lambda}\left(\sum_{k=l}^{\infty} c_k e^{-ik\lambda}\right)$$

$$=e^{il\lambda}\left(\sum_{k=l-1}^{\infty} c_k e^{-ik\lambda}-c_{l-1}e^{-i(l-1)\lambda}\right)$$

$$=e^{i\lambda}\left(e^{i(l-1)\lambda}\sum_{k=l-1}^{\infty} c_k e^{-ik\lambda}-c_{l-1}\right)$$

$$=e^{i\lambda}(W_l(e^{-i\lambda})-c_{l-1}), \tag{2.200}$$

$$l=1,2,\cdots,n-1,$$

因而

$$Y(t)=\int_{\Pi} e^{i\lambda t}\begin{bmatrix} W_1(e^{-i\lambda}) \\ W_2(e^{-i\lambda}) \\ \vdots \\ W_n(e^{-i\lambda}) \end{bmatrix} dZ_{\varepsilon}(\lambda). \tag{2.201}$$

由于 $y_j(s)\in H_x(s)$，而 $\varepsilon(t)\perp H_x(s)$，$s<t$，故(2.194)的性质1是显然的。至于$\boldsymbol{Y}(t)$满足(2.195)可由定理2.15的(2.179)和(2.181)直接看出。(2.190)乃是Wold系数的递推公式(见(2.81)式)。▌

如果$x(t)$不是具有谱密度的平稳列，仍可讨论它是否具可扩

张成马氏(多维)序列的问题。J.L Doob 在[13]文中给出了如下相当一般的结果：

定理2.17 (Doob) $x(t)$是一个N维平稳马氏序列的分量的充分必要条件是：$x(t)$的谱函数可表为

$$F_x(\lambda)=\int_{\Pi}\frac{|P_{N-1}(e^{-i\lambda})|^2}{|Q_N(e^{-i\lambda})|^2}d\lambda+G(\lambda), \qquad (2.202)$$

其中：

(1) $G(\lambda)$是非降的阶梯函数，其跳跃点的个数不多于N个。

(2) $Q_N(e^{-i\lambda})$在$G(\lambda)$的诸跳跃点上为零，$P_{N-1}(e^{-i\lambda})$在这些点上也为零，且P_{N-1}的零点阶数不低于Q_N对应零点的阶数。

(3) $P_{N-1}(z)$,$Q_N(z)$的根模不小于1。

显然，有兴趣的一个问题是：假定已知(2.202)中P_{N-1},Q_N以$G(\lambda)$的有关参数，如何具体将$x(t)$马氏扩张以后的$\boldsymbol{Y}(t)$以及相应转移矩阵$\boldsymbol{A}$及$\boldsymbol{C}$能够用这些参数表出，这一问题的解决可参看[36]。

习　题

1. 设x_t满足$x_t+\varphi_1 x_{t-1}=\sigma\varepsilon_t$，$\sigma>0$，$0<|\varphi_1|<1$。试求：

(a) $\Gamma_x(z)$极大函数及 Wold 系数$\{c_k\}$；

(b) 相关函数$R_x(\tau)$及其谱密度。

2. 已知x_t的相关系数值$\rho_0=1$，$\rho_1=0.5$，$\rho_2=0.3$，$\rho_3=0.2$，$\rho_4=0.1$试求$\varphi_k^{(k)}$，$k=1,2,3,4$四个偏相关值。

3. 设x_t是 ARMA 模型，试证逆转公式

$$\varepsilon_t=\sum_{k=0}^{\infty}d_k x_{t-k}$$

中$\{d_k\}$系数与 Wold 系数$\{c_k\}$之间有如下关系

$$\begin{cases} d_0 = 1/c_0; \\ d_\mu = -d_0 \sum_{k=1}^{\mu} c_k d_{\mu-k}, \quad \mu \geq 1. \end{cases}$$

4. 设 x_t 是 AR(1)序列，则

$$\underset{H_x(t)}{\text{Proj}}(x_{t+1}) = \underset{\mathscr{L}(x_t)}{\text{Proj}}(x_{t+1}).$$

5. 设 x_t 是平稳列，令

$$H_x^{+|-} = \{\zeta;\ \zeta = \underset{H_x(0)}{\text{Proj}}(\xi),\ \xi \in H_x[0, +\infty)\}.$$

试证：若 x_t 是 ARMA 模型，则 $\dim\{H_x^{+|-}\} = m < +\infty$。

事实上这一条件也是充分的（请参看第六章§4）。

6. 设 x_t 是零均值平稳列，已知相关系数 $\rho_0 = 1$，$\rho_1 = \frac{1}{2}$，$\rho_2 = 0$，$\rho_3 = \frac{1}{2}$，试证必存在常数 $a > 0$ 和 β，使得 Pr 1 地有

$$x_t = a x_{t-3} + \underset{L(x_{t-1}, x_{t-2})}{\text{Proj}}[x_t - a x_{t-3}] + \beta$$

成立。

7. (Bartlett 公式)设 x_t 为正态 ARMA 序列，$x_1, x_2, \cdots, x_N$ 为 x_t 的一段观测样本，γ_k 为 x_t 的相关函数，令

$$\hat{\gamma}_k = \hat{\gamma}_{-k} = \frac{1}{N} \sum_{l=1}^{N-k} x_l x_{l+k}, \quad k = 0, 1, 2, \cdots, N-1,$$

$$\tilde{\gamma}_k = \hat{\gamma}_k - \gamma_k.$$

则

$$E(\hat{\gamma}_m - \gamma_m)(\hat{\gamma}_n - \gamma_n) = E(\tilde{\gamma}_m \tilde{\gamma}_n)$$

$$= \frac{1}{N} \sum_{l=-\infty}^{\infty} (\gamma_{l-m}\gamma_{l-n} + \gamma_{l-m}\gamma_{l+n}) + O\left(\frac{1}{N^2}\right).$$

（注：利用如下事实——设 x_i 为 0 均值正态变量，则

$$E x_i x_j x_k x_l = \sigma_{i,j}\sigma_{k,l} + \sigma_{i,k}\sigma_{j,l} + \sigma_{i,l}\sigma_{j,k}.$$ ）

8. 设 x_t 为实 ARMA 正态序列，则

$$\lim_{N\to\infty} N\,E(\hat{\gamma}_k\hat{\gamma}_l)=4\pi\int_{-\pi}^{\pi} f_x^2(\lambda)\cos l\lambda\cos k\lambda d\lambda.$$

9. 设 x_t 为正态 MA(q)型，则估计 $\hat{\gamma}_k$ 的方差

$$D(\hat{\gamma}_k)=\begin{cases}\dfrac{2}{N}\left(\gamma_0^2+2\sum\limits_{l=1}^{q}\gamma_l^2\right)+O\left(\dfrac{1}{N^2}\right), & k=0;\\[2ex] \dfrac{1}{N}\left(\gamma_0^2+2\sum\limits_{l=1}^{q}\gamma_l^2\right)+O\left(\dfrac{1}{N^2}\right), & |k|>q.\end{cases}$$

10. 设 x_t 为平稳列，令

$$H_x(t)=\mathscr{L}\{x_s;s\leqslant t\};$$

$$S_x=\bigcap_t H_x(t);$$

$$H_x=\mathscr{L}\{x_t;t=0,\pm1,\cdots\}.$$

称 x_t 是奇异的，若 $S_x=H_x$。记

$$s_x(t)=\operatorname*{Proj}_{S_x}(x_t),$$

则：

(a) $s_x(t)$仍是平稳列；

(b) $s_x(t)$是奇异的。

11. 设 x_t 是平稳列，令

$$u_x(t)=\frac{\Delta(t)}{\|\Delta(t)\|},$$

其中

$$\Delta(t)=x_t-E[x_t|H_x(t-1)].$$

则 x_t 是非奇异的(即 $H_x\neq H_x(t)$)$\Longleftrightarrow\|\Delta(t)\|\neq0$。

12. 设 x_t 是非奇异的，则 $u_x(t),u_x(t-1),\cdots$组成了 $H_x(t)\ominus S_x$ 的完全标准正交基。

13. 如果是 $x_t\neq0$的平稳列，并 $s_x(t)=0$则称x_t是正则序列，于是 x_t 是正则序列$\Longleftrightarrow S_x=\varnothing$。

11. 设 x_t 是非奇异的平稳列，则 x_t 可唯一地分解为 $x_t = \gamma_t \oplus s_t$，其中 γ_t 是正则序列，s_t 是奇异序列，两者是正交的。

15. 设 x_t 是平稳列，具有 p 步复杂马氏性：

$$\underset{H_x(t)}{\mathrm{Proj}}(x_{t+\tau}) = \underset{L\{x_t, x_{t-1}, \cdots, x_{t-p+1}\}}{\mathrm{Proj}} (x_{t+\tau}) \quad (\tau > 0),$$

则 x_t 必为 AR 序列，其阶数不超过 p。

16. 设 x_t 为 AR(p) 模型，则由(2.121)递推解出的 $\{\varphi_j^{(k)}\}$ 和

$$\theta_0^2(k) = R(0) - \sum_{l=1}^{k} R(l)\varphi_l^{(k)}, \qquad 1 \leqslant k \leqslant p$$

之间有如下关系

$$\theta_0^2(k+1) = \theta_{0,0}^2 \prod_{l=1}^{k+1} (1 - (\varphi_l^{(l)})^2), \qquad 1 \leqslant k < p,$$

而

$$\theta_{0,0}^2 = R(0).$$

第三章 时间序列的预报与滤波

§1 引 言

1° 时间序列的预报问题

在许多自然科学、工程技术、社会经济等问题中，经常会出现"根据以往的历史观测"希望对"将来的某个时刻"的观测作出预报的问题，并且希望预报得"愈精确"愈好。从数学上讲就是：设有时间序列$\cdots,x_{-3},x_{-2},x_{-1},x_0$为已知历史观测数据，希望对未来的$x_\tau(\tau>0)$作最优的预报，也就是要找一个预报函数或称预测函数

$$\varphi_\tau=\varphi_\tau(x_t|t\leqslant 0), \tag{3.1}$$

在某种准则下，比如Z，使得

$$Z(x_\tau,\varphi_\tau) \tag{3.2}$$

为极小值或极大值。于是根据不同的实际问题和不同的背景，可以讨论在不同的φ_τ函数和准则Z之下的时间序列预测问题。比如φ_τ可以研究是$\{x_t|t\leqslant 0\}$的线性函数也可以讨论是非线性函数；关于标准Z，可以取为均方误差标准，也可以选条件概率，也可以选误差概率等等。

以下我们只讨论一种比较容易处理而在实用中最有广泛意义的场合：即φ_τ是线性的，Z为均方误差$E|x_\tau-\varphi_\tau|^2$的准则。至于非线性的φ_τ和非均方误差准则下的讨论，读者可参看[22]，[24]。均方误差准则是工程、技术、物理等实际部门广为流传和通用的，至于我们只限定在线性函数的范围内来寻找最优的φ_τ，除了线性函数比较容易处理和应用这一理由之外，更重要的是因为我们有以下的重要的概率论背景：

定理3.1 设x_t是Gauss平稳列($Ex_t\equiv 0$)，对任何$t_1<t_2<\cdots<t_n<\tau$的整数列，记

$$\hat{L}_\tau(x_{t_1}, x_{t_2}, \cdots, x_{t_n}) \tag{3.3}$$

为利用$\{x_{t_1}, \cdots, x_{t_n}\}$资料对$x_\tau$在均方误差标准下的最优线性预测，记

$$E[x_\tau | \mathscr{F}(x_{t_1}, \cdots, x_{t_n})] \tag{3.4}$$

为x_τ在由$\{x_{t_1}, \cdots, x_{t_n}\}$所产生的$\sigma$代数$\mathscr{F}(x_{t_1}, \cdots, x_n)$下的条件期望，则

$$E[x_\tau | \mathscr{F}(x_{t_1}, \cdots, x_{t_n})] = \hat{L}_\tau(x_{t_1}, \cdots, x_{t_n}). \tag{3.5}$$

此定理的证明可参考[2]，这里不细说了。但是这一定理的(3.5)式告诉我们：当讨论的对象是 Gauss 平稳序列时，最好的预报量，在均方准则下，可由线性预报量达到。而 Gauss 序列又是实用中最有代表性的。因此今后为简明起见，我们只讨论线性预测问题和线性滤波问题。

2° 时间序列的滤波问题

在现实的观测中，绝大多数场合都有各种干扰因素存在于观测系统之中，这类干扰通常称之为“噪声”。我们观测到的数据并不是真正研究对象的测量值而是“信号”与“噪声”的混叠。设以x_t表研究对象的测量真值，n_t为噪声，则观测到的是

$$y_t = x_t + n_t. \tag{3.6}$$

假定对x_t, n_t分别掌握了它们的统计性质，是否能从y_t的观测数据中恢复x_t呢？甚至能够对$x_{t+\tau}$ $(\tau>0)$作出预报呢？简言之，滤波问题主要的是从y_t观测中“滤掉”噪声n_t。

滤波问题初一看和预测问题似乎是两回事，一个是要预报另一个是要滤噪声。然而滤波问题在理论上和预测问题却是紧密相关的。

设$\{x_t;\ t=0, \pm 1, \cdots\}$是平稳列，已知$\{x_s; s\leqslant t\}$求对$x_{t+\tau}$ $(\tau>0)$的最优预测问题，理论上的回答如下：如果最优准则是采用均方标准，预测量是线性的，此时可记

$$H_x(t) = \mathscr{L}\{x_s;\ s\leqslant t\},$$

$$H_x=\mathscr{L}\{x_s;\ s\in\text{整数集}\}.$$

显见 $x_{t+\tau}\in H_x$，$H_x(t)\subset H_x$，则问题变为求 $\zeta_t^{(\tau)}\in H_x(t)$ 使

$$E|x_{t+\tau}-\zeta_t^{(\tau)}|^2=\|x_{t+\tau}-\zeta_t^{(\tau)}\|^2$$
$$=\inf_{\xi\subset H_x(t)}\|x_{t+\tau}-\xi\|^2. \tag{3.7}$$

由 Hilbert 空间中的 Riesz 分解定理(见[5])易知

$$\zeta_t^{(\tau)}=\underset{H_x(t)}{\mathrm{Proj}}(x_{t+\tau})\triangleq E[x_{t+\tau}|H_x(t)]. \tag{3.8}$$

对于滤波问题，假若 x_t, n_t 是平稳序列（一般讲噪声与信号是独立的，故 x_t 与 n_t 独立)，已知 $\{y_s;s\leqslant t\}$ 之下求对 $x_{t+\tau}$ 的滤波问题可用类似于预测方法来解。

记

$$H_y(t)=\mathscr{L}\{y_s;\ s\leqslant t\}, \tag{3.9}$$

$$H=\{\text{一切二阶矩变量组成的希氏空间}\}^{①}.$$

显见

$$H_y(t)\subset H,\quad x_{t+\tau}\in H. \tag{3.10}$$

要求最优滤波 $\zeta_{t+\tau}$，等于求 $\zeta_{t+\tau}$ 使

$$\|x_{t+\tau}-\zeta_{t+\tau}\|^2=\inf_{\xi\in H_y(t)}\|x_{t+\tau}-\xi\|^2. \tag{3.11}$$

显然

$$\zeta_{t+\tau}=\underset{H_y(t)}{\mathrm{Proj}}(x_{t+\tau}). \tag{3.12}$$

为了便于解决问题，往往是先求

$$\eta_\tau=\underset{H_y}{\mathrm{Proj}}(x_{t+\tau}), \tag{3.13}$$

然后再求

$$\zeta_{t+\tau}=\underset{H_y(t)}{\mathrm{Proj}}(\eta_\tau). \tag{3.14}$$

因为 $H_y(t)\subset H_y$，因而两次投影和 $x_{t+\tau}$ 直接对 $H_y(t)$ 的投影是一样的，但 $\eta_\tau\in H_y$，因而

$$\zeta_{t+\tau}=E[\eta_\tau|H_y(t)]. \tag{3.15}$$

---

① 参看 (1.33) 式的有关段落。

当$\tau=0$称为滤波，当$\tau>0$称为预测性滤波，当$\tau<0$称为滞后滤波(这一场合在预测问题是没有的)。

3° 用 Wold 分解来解决预报问题

在第二章中，我们已经知道，设x_t是平稳 ARMA 序列，则有 Wold 分解式

$$x_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k},\qquad t=0,\pm 1,\cdots. \tag{3.16}$$

而$\{\varepsilon_t;t=0,\pm 1,\cdots\}$构成了$H_x$中的标准正交基，并且具有极其重要的等时从属性质：

$$\varepsilon_t\in H_x(t)\ominus H_x(t-1) \tag{3.17}$$

(参看定理2.5)。因此由(3.16)有

$$\begin{aligned}\underset{H_x(t)}{\mathrm{Proj}}\,(x_{t+\tau})&=\sum_{k=0}^{\infty}c_k\underset{H_x(t)}{\mathrm{Proj}}(\varepsilon_{t+\tau-k})\\&=\sum_{k=\tau}^{\infty}c_k\varepsilon_{t-\tau-k}.\end{aligned} \tag{3.18}$$

当x_t的 ARMA 模型给定之后，$\{c_k\}$可以由模型参数递推出来(参看(2.81)式)，而ε_t是

$$\Phi(U)x_t=\Theta(U)\varepsilon_t$$

中的“新息”(见(2.72)式)。根据(2.82)及(3.18)可得

$$\underset{H_x(t)}{\mathrm{Proj}}\,(x_{t+\tau})=\hat{L}(x_s\,|\,s\leqslant t), \tag{3.19}$$

其中$\hat{L}$中参数由$\{c_k,d_s\}$确定，从而原则上回答了预报的理论问题。

至于滤波问题，在求得η_τ之后(见(3.13))，因$\eta_\tau\in H_y$，故由$\{\varepsilon_t^{(y)}\}$Wold 序列，可正交展开得

$$\eta_\tau=\sum_{k=-\infty}^{\infty}\beta_{k+\tau}^{(\eta)}\varepsilon_t^{(y)}. \tag{3.20}$$

于是类似于(3.18)有

$$\zeta_{t+\tau}=E[\eta_\tau|H_y(t)]=\sum_{k=-\infty}^{t}\beta_{k,\tau}^{(\eta)}\varepsilon_k^{(y)}. \qquad (3.21)$$

从而理论上也回答了如何求最优滤波的问题。

但是，以上的求法在应用中还是很繁难的，针对ARMA模型许多学者提出了一系列的预测和滤波方法，简单归纳可分为频率域的方法和时间域的方法，前者以苏联数学家Яглом为代表（见[11]），后者以美国的Box，Jenkins和Kalman（见[25]，[26]，[27]）为代表。这两类方法各有优点，我们在以下各节中就其方法的实质各作介绍。

§2 ARMA模型的预测方法

1° ARMA序列预测的频域方法

以下将介绍的是具有非常实用价值的Яглом方法。所谓“频域”方法是指：对ARMA序列x_t，我们已经得到了它的有理谱密度

$$f_x(\lambda)=\frac{1}{2\pi}\left|\frac{\sum_{k=0}^{q}\theta_k e^{-ik\lambda}}{\sum_{l=0}^{p}\varphi_l e^{il\lambda}}\right|^2, \qquad -\pi\leqslant\lambda\leqslant\pi \qquad (3.22)$$

（见(2.50)式），以此为出发点，在获得观测资料$\{x_s; s\leqslant t\}$后，给出均方标准下的最优预报$\hat{x}_{t+\tau}$。

先证以下二条关键性的定理：

定理3.2 设x_t为平稳列，谱密度为$f_x(\lambda)$，在已知$\{x_\tau; \tau<0\}$条件下，对x_M（$M\geqslant0$）的最优预报公式为

$$\hat{x}_M=\int_{-\pi}^{\pi}\varphi_M(\lambda)dZ_x(\lambda), \qquad (3.23)$$

其中复函数$\varphi_M(\lambda)=\Phi(z)\Big|_{z=e^{i\lambda}}$，而$\Phi(z)$由以下确定：

(1) $\Phi_M(z)$在$|z|>1-\theta$ $(\theta>0)$上解析，并且

$$\Phi_M(\infty)=0. \tag{3.24}$$

(2) $[z^M-\Phi_M(z)]f_x^*(z)$ 在$|z|<1+\rho$ $(\rho>0)$内解析，其中$f_x^*(z)$是复函数使

$$f_x^*(e^{i\lambda})=f_x(\lambda), \quad -\pi\leqslant\lambda\leqslant\pi. \tag{3.25}$$

证明 由本章引言论述中可看出

$$\hat{x}_M=E[x_M|H_x(-1)]\in H_x(-1). \tag{3.26}$$

然而，由第一章的引理1.1的论述知

$$H_x\longleftrightarrow L^2(dF_x) \quad \text{(同构对应)},$$

$$H_x(t)\longleftrightarrow L_t^2(dF_x) \quad \text{(同构对应)}. \tag{3.27}$$

其中$L_t^2(dF_x)=\mathscr{L}\{e^{ik\lambda};k\leqslant t\}$(依$\|\ \ \|^2dF_x$取线性闭包)，因此(3.26)正交投影决定于

$$(x_M-\hat{x}_M,x_k)=0, \quad k=-1,-2,\cdots, \tag{3.28}$$

对应于函数空间中的

$$\int_{-\pi}^{\pi}(e^{iM\lambda}-\varphi_M(\lambda)(\overline{e^{ik\lambda}})dF_x(\lambda)=0, \quad k=-1,-2,\cdots, \tag{3.29}$$

即

$$\int_{-\pi}^{\pi}(e^{iM\lambda}-\varphi_M(\lambda)e^{ik\lambda}f_x(\lambda)d\lambda=0, \quad k=1,2,\cdots. \tag{3.30}$$

由定理的条件1，

$$\Phi_M(z)=\sum_{k=1}^{\infty}\frac{a_k}{z^k} \tag{3.31}$$

(k从1开始是因为$\Phi_M(\infty)=0$，不可能有常数项)，故

$$\Phi_M(e^{i\lambda})=\sum_{k=1}^{\infty}a_ke^{-ik\lambda}. \tag{3.32}$$

又由条件2，则

$$[z^M-\Phi_M(z)]f_x^*(z)=\sum_{k=0}^{\infty}b_kz^k$$

在比1大一点的圆域内一致收敛，故

$$[e^{iM\lambda}-\Phi_M(e^{i\lambda})]f_x(\lambda)=\sum_{k=0}^{\infty}b_k e^{ik\lambda}. \tag{3.33}$$

可见：(1) $\varphi_M(\lambda)=\Phi_M(e^{i\lambda})\in L^2_{-1}(dF_x)$(见(3.32))。

(2) (3.30)式对应于

$$\int_{-\pi}^{\pi}(e^{iM\lambda}-\varphi_M(\lambda))e^{il\lambda}f_x(\lambda)d\lambda$$

$$=\sum_{k=0}^{\infty}b_k\int_{-\pi}^{\pi}e^{i(k+l)\lambda}d\lambda\equiv 0,\quad 一切\ l>0. \tag{3.34}$$

可见(3.30)满足

$$\hat{x}_M=\int_{-\pi}^{\pi}\varphi_M(\lambda)dz_x(\lambda)=E[x_M|H_x(-1)]. \quad ▌$$

定理3.3 设 x_t 是平稳列，其谱密度为

$$f_x(\lambda)=f^*(e^{i\lambda}),$$

$$f^*(z)=\left|\frac{B_0}{A_0}\right|^2\frac{(z-b_1)\cdots(z-b_q)(1-\bar{b}_1z)\cdots(1-\bar{b}_qz)}{(z-a_1)\cdots(z-a_p)(1-\bar{a}_1z)\cdots(1-\bar{a}_pz)}z^{p-q}, \tag{3.35}$$

其中

$$0<|a_i|<1,\quad 0<|b_j|<1.$$

当 $p\geqslant q$ 时，确定一个 $p-1$ 次的多项式 $Q(z)$ 使得

$$z^{M+p-q}(z-b_1)\cdots(z-b_q)-Q(z) \tag{3.36}$$

以 $a_1,a_2,\cdots,a_p$ 为根①；

当 $p<q$ 时，确定一个 $q-1$ 次的多项式 $Q(z)$ 使得

$$z^M(z-b_1)\cdots(z-b_q)-Q(z) \tag{3.37}$$

以 $a_1,a_2,\cdots,a_p$ 为根并具有 $q-p$ 个零根。

令

① 可能还有别的根。

$$\Phi_M(z)=\begin{cases}\dfrac{Q(z)}{z^{p-q}(z-b_1)\cdots(z-b_q)}, & p\geqslant q;\\[2ex] \dfrac{Q(z)}{(z-b_1)(z-b_2)\cdots(z-b_q)}, & p<q.\end{cases} \tag{3.38}$$

则已知$\{x_s;s<0\}$之下对x_M，$M\geqslant 0$的最优预报为

$$\hat{x}_M=\int_{-\pi}^{\pi}\Phi_M(e^{i\lambda})dZ_x(\lambda), \tag{3.39}$$

并且记预测误差为σ_M^2，则

$$\lim_{M\to\infty}\|x_M-\hat{x}_M\|^2=\lim_{M\to\infty}\sigma_M^2$$

$$=\int_{-\pi}^{\pi}f(\lambda)d\lambda=R_x(0). \tag{3.40}$$

证明 当$p\geqslant q$时，$Q(z)$是$p-1$次多项式，

$$\Phi_M(z)=\frac{Q(z)}{z^{p-q}(z-b_1)(z_2-b_2)\cdots(z-b_q)}$$

是真分式(分母阶数为p)，故$\Phi_M(\infty)=0$，并且容易看出$\Phi_M(z)$在$|z|>1-\epsilon$上是解析的(因为$|b_j|<1$)。又

$$z^M-\Phi_M(z)=z^M-\frac{Q(z)}{z^{p-q}(z-b_1)\cdots(z-b_q)}$$

$$=\frac{z^{M+p-q}(z-b_1)\cdots(z-b_q)-Q(z)}{z^{p-q}(z-b_1)\cdots(z-b_q)}, \tag{3.41}$$

若要求(3.41)的分子式以$a_1,a_2,\cdots,a_p$为根，则应可表为

$$z^M-\Phi_M(z)=\frac{(z-a_1)\cdots(z-a_p)\gamma(z)}{z^{p-q}(z-b_1)\cdots(z-b_q)} \tag{3.42}$$

的形式，$\gamma(z)$在全平面解析。由假设(3.35)

$$(z^M-\Phi_M(z))f^*(z)=\left|\frac{B_0}{A_0}\right|^2\frac{(1-\bar{b}_1z)(1-\bar{b}_2z)\cdots(1-\bar{b}_qz)}{(1-\bar{a}_1z)(1-\bar{a}_2z)\cdots(1-\bar{a}_pz)}\gamma(z) \tag{3.43}$$

在$|z|<1+\rho(\rho>0)$内是解析的(因为$0<|a_i|<1,0<|b_j|<1$)。由于$\Phi_M(z)$满足定理3.2的条件，则

$$\hat{x}_M = \int_{-\pi}^{\pi} \varphi_M(\lambda) \mathrm{d} z_x(\lambda)$$

是最优预报，其中 $\varphi_M(\lambda) = \Phi_M(\mathrm{e}^{\mathrm{i}\lambda})$。

对于 $p<q$，用以上类似的方法可证明本定理的结论。

此外，由 Wold 分解知

$$\sigma_M^2 = \|x_M - \hat{x}_M\|^2 = \sum_{k=0}^{M} |c_k|^2.$$

$$\lim_{M\to\infty} \sigma_M^2 = \sum_{k=0}^{\infty} |c_k|^2 = \|x_t\|^2 = R_x(0) = \int_{-\pi}^{\pi} f_x(\lambda) \mathrm{d}\lambda. \quad \blacksquare$$

以上介绍的方法，对已知 ARMA 模型谱密度的条件下是很实用的，以后常称 $\varphi_M(\lambda)$ 为 M 步预报的谱特征。

例3.1 设 x_t 是AR(1)序列，其谱密度为

$$f_x(\lambda) = \frac{B_0^2}{|\mathrm{e}^{-\mathrm{i}\lambda} - \varphi_1|^2}, \quad 1 < |\varphi_1|. \qquad (3.44)$$

求在已知$\{x_s; s<0\}$条件下对 x_M 的预报公式

解 先变形使具(3.35)的形式

$$f_x(\lambda) = \frac{B_0^2}{|\varphi_1|^2 \left| \mathrm{e}^{\mathrm{i}\lambda} - \frac{1}{\varphi_1} \right|^2} \quad (p=1, q=0),$$

则

$$f^*(z) = \left|\frac{B_0}{\varphi_1}\right|^2 \frac{1}{\left(z - \frac{1}{\varphi_1}\right)\left(1 - \frac{1}{\varphi_1} z\right)} z,$$

$$f^*(\mathrm{e}^{\mathrm{i}\lambda}) = \left|\frac{B_0}{\varphi_1}\right|^2 \frac{1}{\left(\mathrm{e}^{\mathrm{i}\lambda} - \frac{1}{\varphi_1}\right)\left(\mathrm{e}^{-\mathrm{i}\lambda} - \frac{1}{\varphi_1}\right)} = f_x(\lambda).$$

由 $p>q$，则选 $Q(z)$ 为 $p-1=0$ 次多项式——常数 c 使

$$z^M z^{p-q}(z-b_1)\cdots(z-b_q) - Q(z) = z^{M+1} - c$$

以$\frac{1}{\varphi_1}$为根，由此知

$$c = \varphi_1^{-(M+1)}. \tag{3.45}$$

可见M步预报的谱特征为

$$\varphi_M(\lambda) = \Phi_M(z)\big|_{z=e^{i\lambda}} = \frac{Q(z)}{z^p}\bigg|_{z=e^{i\lambda}}$$

$$= \frac{\varphi_1^{-(M+1)}}{z}\bigg|_{z=e^{i\lambda}} = \frac{e^{-i\lambda}}{\varphi_1^{(M+1)}}. \tag{3.46}$$

从而

$$\hat{x}_M = \int_{-\pi}^{\pi} \Phi_M(e^{i\lambda}) dZ_x(\lambda)$$

$$= \frac{1}{(\varphi_1)^{M+1}} \int_{-\pi}^{\pi} e^{-i\lambda} dZ_x(\lambda) = \frac{x_{-1}}{(\varphi_1)^{M+1}}. \tag{3.47}$$

(3.47)的结果我们不会感到惊奇，因为具有(3.44)谱密度的序列是AR(1)序列，因而是马氏型序列，显然应有

$$\hat{x}_M = E[x_M | H_x(-1)] = E[x_M | \mathscr{L}\{x_{-1}\}] = cx_{-1}.$$

例3.2 设x_t的谱密度为

$$f_x(\lambda) = c|e^{i\lambda} - b|^2, \quad 0 < |b| < 1. \tag{3.48}$$

试求M步预报公式。

解 显见，此时$p=0, q=1$ $(p<q)$。

$$f^*(z) = c\frac{(z-b)(1-bz)}{1}z^{-1}.$$

因此，应找$Q(z)$($q-1$阶多项式)即常数a使得

$$z^M(z-b) - Q(z)$$

以$z=0$为根(因$p=0, q=1$)，于是：当$M=0$时，

$(z-b)|_{z=0} - a = 0$，即$a=-b$，

故 $\Phi_M(z) = -b/(z-b) = \frac{(-b)}{z}\sum_{k=0}^{\infty}\left(\frac{b}{z}\right)^k, \quad |z|>1.$

从而

$$
\begin{aligned}
\hat{x}_M &= \int_{-\pi}^{\pi} \Phi_M(e^{i\lambda}) dZ_x(\lambda) \\
&= -\sum_{k=1}^{\infty} \int_{-\pi}^{\pi} e^{-ik\lambda} b^k dZ_x(\lambda) \\
&= -\sum_{k=1}^{\infty} b^k x_{-k}.
\end{aligned}
$$

当$M>0$时：$z^M(z-b)-a=0$ 要有零根只有$a=0$。即这种场合唯有 $\Phi_M(z)=0$，即

$$
\hat{x}_M \equiv 0. \tag{3.49}
$$

显然对(3.49)的结果，从理论上看也是容易理解的，因为具有(3.48)型谱密度的平稳列实质上是MA(1)，其相关函数是一步截尾的，也就是说间隔 2 步以上的随机变量之间是无关的，因此$x_M(M>0)$与 $H_x(-1)$的任一元素无关，故预报量为零。

例3.3 设 x_t 平稳列具有谱密度

$$
f_x(\lambda) = \frac{c}{|e^{i\lambda}-a_1|^2 |e^{i\lambda}-a_2|^2} \quad (0<|a_i|<1).
$$

求已知$\{x_s, s<0\}$之下对x_M $(M>0)$的预报谱特征。

解 由 $f_x(\lambda)$知 $p=2, q=0$。

$$
f^*(z) = \frac{cz^2}{(z-a_1)(z-a_2)(1-\bar{a}_1 z)(1-\bar{a}_2 z)}.
$$

令（以下换一种等价解法，可参看(3.41)和(3.43)）

$$
\Phi_M(z) = \frac{\alpha z+\beta}{z^2}.
$$

要求 $z=a_1, a_2$ 时为 $z^M - \Phi_M(z)$的根，于是

$$
\begin{cases}
a_1^M - \dfrac{\alpha a_1+\beta}{a_1^2} = 0, \\[2ex]
a_2^M - \dfrac{\alpha a_2+\beta}{a_2^2} = 0.
\end{cases} \tag{3.50}
$$

解(3.50)可得

$$\begin{cases} \alpha = \dfrac{a_1^{M+2} - a_2^{M+2}}{a_1 - a_2}, \\ \beta = \dfrac{a_1 a_2 (a_2^{M+1} - a_1^{M+1})}{a_1 - a_2}. \end{cases} \tag{3.51}$$

于是

$$\varphi_M(\lambda) = \Phi_M(e^{i\lambda}) = \frac{a_1^{M+2} - a_2^{M+2}}{a_1 - a_2} e^{-i\lambda} + \frac{a_1 a_2 (a_2^{M+1} - a_1^{M+1})}{a_1 - a_2} e^{-i2\lambda}.$$

即

$$\hat{x}_M = \frac{a_1^{M+2} - a_2^{M+2}}{a_1 - a_2} x_{-1} + \frac{a_1 a_2 (a_2^{M+1} - a_1^{M+1})}{a_1 - a_2} x_{-2}. \tag{3.52}$$

2° ARMA序列预测的时域方法

所谓时域方法指的是不以谱密度为已知的出发点而由模型和观测数据为讨论问题的出发点。以下根据AR,MA,ARMA三种不同模型建立不同的预报公式。

(1) AR序列的预报公式

设 x_t 为 AR 序列，其模型方程为

$$x_t + \sum_{k=1}^{p} \varphi_k x_{t-k} = \theta_0 \varepsilon_t, \tag{3.53}$$

则

$$x_{t+\tau} = -\sum_{k=1}^{p} \varphi_k x_{t+\tau-k} + \theta_0 \varepsilon_{t+\tau}, \quad \tau > 0. \tag{3.54}$$

将(3.54)对 $H_x(t)$投影可得

$$E[x_{t+\tau} | H_x(t)]$$

$$= -\sum_{k=1}^{p} \varphi_k E[x_{t+\tau-k} | H_x(t)] + \theta_0 E[\varepsilon_{t+\tau} | H_x(t)].$$

然而 $\varepsilon_{t+\tau} \perp H_x(t), \tau > 0$，故

$$E[x_{t+\tau} | H_x(t)] = -\sum_{k=1}^{p} \varphi_k E[x_{t+\tau-k} | H_x(t)]. \tag{3.55}$$

$\tau=1$,

$$\hat{x}_{t+1}=-\sum_{k=1}^{p}\varphi_k x_{t+1-k}. \tag{3.56}$$

$\tau=2$,

$$\begin{aligned}\hat{x}_{t+2}&=-\left(\sum_{k=2}^{p}\varphi_k x_{t+2-k}+\varphi_1 E[x_{t+1}|H_x(t)]\right)\\&=-\sum_{k'=0}^{p-2}\varphi_{k'+2}x_{t-k'}+\varphi_1\sum_{k=1}^{p}\varphi_k x_{t+1-k}\\&=-\sum_{k'=0}^{p-2}(\varphi_{k'+2}-\varphi_1\varphi_{k'+1})x_{t-k'}+\varphi_1\varphi_p x_{t-p+1}.\end{aligned} \tag{3.57}$$

由(3.55),(3.56),(3.57)可知:

(a) 如果已知$\{x_s;\ t-p<s\leqslant t\}$及$\{E[x_{t+\tau-1}|H_x(t)]\}$有关预报量，则(3.55)告诉我们可以递推出下一步的预报公式。

(b) 由于(3.55)中的$\{E[x_{t+\tau-k}|H_x(t)];\ k=1,\cdots,p\}$皆为$H_x(t)$中的元素，表明AR的预报公式只含“历史”的有限个资料，(3.57)则进一步指明可用有限个“过去”的观测$\{x_{t-l};\ l=0,1,\cdots,t-p+1\}$来明显表示预测量。当然对一般的$\tau$还应有证明，这就是

定理3.4 设x_t为AR(p)模型($p\geqslant1$)，则对$\tau>0$，有

$$E[x_{t+\tau}|H_x(t)]=\frac{1}{\theta_0}\sum_{j=0}^{p-1}\beta_j^{(\tau)}x(t-j), \tag{3.58}$$

其中

$$\begin{bmatrix}\beta_0^{(\tau)}\\\beta_1^{(\tau)}\\\vdots\\\beta_{p-1}^{(\tau)}\end{bmatrix}=\begin{bmatrix}1 & & & 0 & \\ \varphi_1 & 1 & & & \\ & & \ddots & & \\ \varphi_{p-1} & \varphi_{p-2} & \cdots & \varphi_1 & 1\end{bmatrix}\begin{bmatrix}c_\tau\\c_{\tau+1}\\\vdots\\c_{\tau+p-1}\end{bmatrix}, \tag{3.59}$$

并且

$$c_s=\begin{cases}\theta_0, & s=0;\\ -\sum_{l=1}^{s}\varphi_l c_{s-l}, & 1\leqslant s\leqslant p;\\ -\sum_{l=1}^{p}\varphi_l c_{s-l}, & p<s.\end{cases}\tag{3.60}$$

证明　由 AR 模型的 Wold 分解知

$$x_{t+\tau}=\sum_{k=\tau}^{\infty}c_k\varepsilon_{t+\tau-k}.$$

$$E[x_{t+\tau}|H_x(t)]=\sum_{k=\tau}^{\infty}c_k\varepsilon_{t+\tau-k}$$

$$=\int_{-\pi}^{\pi}e^{i(t+\tau)\lambda}\left(\sum_{k=\tau}^{\infty}c_k e^{-ik\lambda}\right)\Gamma_x^{-1}(e^{-i\lambda})dZ_x(\lambda).\tag{3.61}$$

其中 $\Gamma_x(z)$ 为 AR 模型的极大解析函数(见(2.41))，故

$$E[x_{t+\tau}|H_x(t)]$$

$$=\frac{1}{\theta_0}\int_{-\pi}^{\pi}e^{i(t+\tau)\lambda}\left(\sum_{k=\tau}^{\infty}\sum_{l=0}^{p}c_k\varphi_l e^{-i(k+l)\lambda}\right)dZ_x(\lambda).\tag{3.62}$$

将(3.62)求和部分交换求和次序可得

$$V\triangleq\sum_{k=\tau}^{\infty}\sum_{l=0}^{p}c_k\varphi_l e^{-i(k+l)\lambda}$$

$$=\left(\sum_{s=\tau}^{\tau+p}\sum_{l=0}^{s-\tau}+\sum_{s=\tau+p+1}^{\infty}\sum_{l=0}^{p}\right)e^{-is\lambda}c_{s-l}\varphi_l.\tag{3.63}$$

然而由 Wold 系数的递推公式(2.81)知(3.60)成立，因此(3.63)中的第二求和项

$$\sum_{s=\tau+p+1}^{\infty}\left(\sum_{l=0}^{p}\varphi_l c_{s-l}\right)\equiv 0\ (s>p).$$

故

$$V=\left(\sum_{s=\tau}^{\tau+p}\sum_{l=0}^{s-\tau}\varphi_l c_{s-l}\right)e^{-is\lambda}$$

$$\xlongequal{j=s-\tau}\sum_{j=0}^{p}\left(\sum_{l=0}^{j}\varphi_l c_{j+\tau-l}\right)e^{-i\lambda(j+\tau)}. \tag{3.64}$$

但是当 $j=p$ 时(3.64)的括号项为零，故只有 p 项，代回(3.62)式

$$E[x_{t+\tau}|H_x(t)]=\frac{1}{\theta_0}\int_{-\pi}^{\pi}e^{i\lambda(t+\tau)}\left(\sum_{j=0}^{p-1}\beta_j^{(\tau)}e^{-i\lambda(j+\tau)}\right)dZ_x(\lambda)$$

$$=\sum_{j=0}^{p-1}\frac{1}{\theta_0}\beta_j^{(\tau)}\int_{-\pi}^{\pi}e^{i\lambda(t-j)}dZ_x(\lambda)$$

$$=\sum_{j=0}^{p-1}\frac{1}{\theta_0}\beta_j^{(\tau)}x_{t-j}, \tag{3.65}$$

其中

$$\begin{cases}\beta_0^{(\tau)}=c_\tau;\\ \beta_j^{(\tau)}=\sum_{l=0}^{j}\varphi_l c_{j+\tau-l}, \quad 0<j\leqslant p-1.\end{cases} \tag{3.66}$$ ▌

例3.4 设某地区年平均降水量为540(mm)，其偏差量 x_t 满足以下 AR(2)方程

$$x_t+0.54x_{t-1}-0.3x_{t-2}=\varepsilon_t.$$

近五年的实际降水量为

…	y_{-4}	y_{-3}	y_{-2}	y_{-1}	y_0	$\bar{y}=540$
…	560	470	585	496	576	$y_t=\bar{y}+x_t$

求预测今后 3 年内的降水量。

解 对 $\tau=1$，利用(3.56)得

$$x_{t+1}=-0.54x_t+0.3x_{t-1},$$

即

$$\hat{x}_1=-0.54(576-540)+0.3(496-540)=-32.64.$$

故

$$\hat{y}_1 = \hat{x}_1 + \bar{y} = 540 - 32.64 = 507.36.$$

计算机模型模拟为508。

对 $\tau = 2$，利用(3.56)得

$$\hat{x}_2 = -\varphi_2 x_0 - \varphi_1 \hat{x}_1$$
$$= +0.3(576-540) - 0.54 \times (-32.64) = 28.43.$$

故

$$\hat{y}_2 = \hat{x}_2 + \bar{y} = 540 + 28.43 = 568.43 \text{（模拟值为567）}.$$

对 $\tau = 3$，用递推公式

$$E[x_3 | H_x(0)] = -\varphi_1 \hat{x}_2 - \varphi_2 \hat{x}_1$$
$$= -0.54 \times 28.43 - (-0.3) \times (-32.64)$$
$$= -25.14.$$

故

$$\hat{y}_3 = \hat{x}_3 + \bar{y} = 540 - 25.14 = 514.86.$$

对 $\tau = 3$，可直接用(3.58)的公式：

先求$\{c_k\}$。本题

$$\theta_0 = 1,\ \varphi_1 = 0.54,\ \varphi_2 = -0.3,\ p = 2.$$

故

$$c_0 = \theta_0 = 1, c_1 = -\varphi_1 c_0 = -0.54, c_2 = -\varphi_1 c_1 - \varphi_2 c_0 = 0.5916,$$
$$c_3 = -\varphi_1 c_2 - \varphi_2 c_1 = -0.4815,\quad c_4 = -\varphi_1 c_3 - \varphi_2 c_2 = 0.4375.$$

故由(3.59)得

$$\begin{bmatrix} \beta_0^{(3)} \\ \beta_1^{(3)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.54 & 1 \end{bmatrix} \begin{bmatrix} c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -0.4815 \\ 0.1775 \end{bmatrix}.$$

$$\hat{x}_3 = -0.48146 x_0 + 0.1775 x_1 = -25.14,\quad \hat{y}_3 = 514.86.$$

(2) 关于 MA(q)序列的预测公式

首先我们回想例3.2，MA(1)当 $M>0$ 即导致预测值 $\hat{x}_M \equiv 0$。事实上对一般的 MA(q)，利用相关函数的 q 步截尾性必导出 $q+1$ 步以上的预测都是平凡的，即

$$E[x_{t+n} | H_x(t)] \equiv 0,\quad n > q.$$

因此我们的目标只要集中在 $1\sim q$ 之间的预测方法就可以了。以下用预报向量递推方法来建立，由

$$\{E[x_{t+\tau}|H_x(t)],\quad 0<\tau\leqslant q\}$$

诸预报值和新获得的数据 x_{t+1}，通过递推公式，求出

$$\{E[x_{t+\tau+1}|H_x(t+1)],\quad 0<\tau\leqslant q\}。$$

以下记

$$\hat{W}_t^q=(\hat{x}_{t,1},\hat{x}_{t,2},\cdots,\hat{x}_{t,q})',\tag{3.67}$$

其中

$$\hat{x}_{t,l}=\hat{x}_{t+l}=E[x_{t+l}|H_x(t)],\quad 0<l\leqslant q,\tag{3.68}$$

即基于 $H_x(t)$ 向前预报 $t+l$ 步。以下称 $\hat{W}_t^q$ 为立足于 t 时刻 MA(q)的预报向量。

对于 MA(q)序列，由第二章的知识知：

设模型为 $x_t=\sum\limits_{k=0}^{q}\theta_k\varepsilon_{t-k}$，则 Wold 系数为

$$\begin{cases}c_0=\theta_0;\quad c_k=\theta_k, & 1\leqslant k\leqslant q;\\ c_k\equiv 0, & k>q。\end{cases}\tag{3.69}$$

故

$$\begin{aligned}E[x_{t+1+\tau}|H_x(t+1)]&\triangleq\hat{x}_{t+1,\tau}=\sum_{k=\tau}^{q}c_k\varepsilon_{t+1+\tau-k},\\ E[x_{t+1+\tau}|H_x(t)]&\triangleq\hat{x}_{t,\tau+1}=\sum_{k=\tau+1}^{q}c_k\varepsilon_{t+\tau+1-k}。\end{aligned}\tag{3.70}$$

可见

$$\hat{x}_{t+1,\tau}=c_\tau\varepsilon_{t+1}+\hat{x}_{t,\tau+1},\quad 0<\tau\leqslant q。\tag{3.71}$$

然而

$$c_0\varepsilon_{t+1}=x_{t+1}-\sum_{l=1}^{q}c_l\varepsilon_{t+1-l},\tag{3.72}$$

并且

$$E[x_{t+1}|H_x(t)]=\sum_{l=1}^{q}c_l\varepsilon_{t+1-l},$$

代入(3.72)可得

$$\varepsilon_{t+1} = \frac{1}{c_0}(x_{t+1} - E[x_{t+1}|H_x(t)])$$

$$= \frac{1}{c_0}(x_{t+1} - \hat{x}_{t,1}). \tag{3.73}$$

故将(3.73)代回(3.71)的右边第一项可得

$$\begin{cases} \hat{x}_{t+1,\tau} = \hat{x}_{t,\tau+1} + \dfrac{c_\tau}{c_0}(x_{t+1} - \hat{x}_{t,1}), & 0<\tau<q; \\ \hat{x}_{t+1,\tau} = \dfrac{c_\tau}{c_0}(x_{t+1} - \hat{x}_{t,1}), & \tau = q. \end{cases} \tag{3.74}$$

(3.74)最后一个等式是因为 $\hat{x}_{t,q+1}=0$。

将(3.74)写成矩阵形式可得

$$\begin{bmatrix} \hat{x}_{t+1,1} \\ \hat{x}_{t+1,2} \\ \vdots \\ \\ \hat{x}_{t+1,q} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \cdots & 0 \\ & & \ddots & \ddots & \\ & \mathbf{0} & & \ddots & 1 \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{t,1} \\ \hat{x}_{t,2} \\ \vdots \\ \hat{x}_{t,q-1} \\ \hat{x}_{t,q} \end{bmatrix} + \begin{bmatrix} \dfrac{c_1}{c_0}(x_{t+1} - \hat{x}_{t,1}) \\ \dfrac{c_2}{c_0}(x_{t+1} - \hat{x}_{t,1}) \\ \vdots \\ \\ \dfrac{c_q}{c_0}(x_{t+1} - \hat{x}_{t,1}) \end{bmatrix}$$

$$= \begin{bmatrix} -\dfrac{c_1}{c_0} & 1 & 0 & \cdots & 0 \\ -\dfrac{c_2}{c_0} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\dfrac{c_{q-1}}{c_0} & 0 & & \cdots & 1 \\ -\dfrac{c_q}{c_0} & 0 & & \cdots & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{t,1} \\ \hat{x}_{t,2} \\ \vdots \\ \hat{x}_{t,q-1} \\ \hat{x}_{t,q} \end{bmatrix}$$

$$+\begin{bmatrix} c_1/c_0 \\ c_2/c_0 \\ \vdots \\ c_{q-1}/c_0 \\ c_q/c_0 \end{bmatrix} x_{t+1}. \tag{3.75}$$

因此可将以上结果归纳成以下定理：

定理3.5 设 x_t 为 MA(q)序列，则有以下向量递推预报公式

$$\hat{\boldsymbol{W}}^q_{t+1} = \boldsymbol{F}\hat{\boldsymbol{W}}^q_t + \boldsymbol{C}x_{t+1}, \tag{3.76}$$

其中

$$\hat{\boldsymbol{W}}^q_{t+i} = (\hat{x}_{t+i,1}, \hat{x}_{t+i,2}, \cdots, \hat{x}_{t+i,q})', \quad i=0,1, \tag{3.77}$$

$$\boldsymbol{F} = \begin{bmatrix} -\frac{c_1}{c_0} & 1 & & \boldsymbol{0} \\ -\frac{c_2}{c_0} & 0 & 1 & \\ \vdots & \vdots & & \ddots \quad 1 \\ -\frac{c_q}{c_0} & 0 & \cdots & 0 \end{bmatrix}, \tag{3.78}$$

$$\boldsymbol{C} = (c_1/c_0, c_2/c_0, \cdots, c_q/c_0)'. \tag{3.79}$$

应用定理3.5于实际问题时，一开始需有 q 个初始预测值组成 $\hat{\boldsymbol{W}}^q_0$，在没有这一初始预测向量时常常就取一组观测数据

$$\hat{\boldsymbol{W}}^q_0 = (x_1, x_2, \cdots, x_q)' \tag{3.80}$$

代替，然后利用 x_{q+1} 用(3.76)递推 $\hat{\boldsymbol{W}}^q_1, \hat{\boldsymbol{W}}^q_2, \cdots$ 直至 $\hat{\boldsymbol{W}}^q_t$，然后加上 x_{t+1} 向前预报 $\hat{\boldsymbol{W}}^q_{t+1}$。为了避免初值 $\hat{\boldsymbol{W}}^q_0$ 的影响，t 应适当地远。

例3.5 已知序列的均值 $\bar{z}=9.95$，z_t 的观测值和 $x_t=z_t-\bar{z}$ 各值如下①：

t	1	2	3	4	5	6	7	8
z_t	10.5	10.1	8.8	9.9	11.3	12.2	11.3	9.8
x_t	0.55	0.15	−1.15	−0.05	1.35	2.25	1.35	−0.15

9	10	11	12	13	14	15	16	17	18	19
9.7	9.5	10	8.9	8.2	10.2	8.8	8.4	9.6	10.2	10.6
−0.25	−0.45	0.05	−1.05	−1.75	0.25	−1.15	−1.55	−0.35	0.25	0.65

由数据进行模型拟合得 x_t 为 MA(1)模型

$$x_t=\varepsilon_t+0.72\varepsilon_{t-1}. \tag{3.81}$$

求：已知$\{z_s;s\leqslant 16\}$对 z_{17}；已知$\{z_s;s\leqslant 17\}$对 z_{18} 及已知$\{z_s;s\leqslant 18\}$对 z_{19} 的诸预报值。

解 由 MA(1)模型知，只能作 $q=1$ 的预报，不必要作二步以上的预报。$c_0=1, c_1=0.72$，由(3.76)知预报公式为

$$\hat{x}_{t+1,1}=0.72(x_{t+1}-\hat{x}_{t,1}). \tag{3.82}$$

由 $t=0$ 开始，取 $\hat{x}_{0,1}=0$，则用(3.82)可计算得以下各预测值：

$\hat{x}_{1,1}=0.396$，$\hat{x}_{2,1}=-0.177$，$\hat{x}_{3,1}=-0.7$，$\hat{x}_{4,1}=0.468$，

$\hat{x}_{5,1}=0.635$，$\hat{x}_{6,1}=1.162$，$\hat{x}_{7,1}=0.135$，$\hat{x}_{8,1}=-0.2052$，

$\hat{x}_{9,1}=-0.0322$，$\hat{x}_{10,1}=-0.300$，$\hat{x}_{11,1}=0.252$，

$\hat{x}_{12,1}=-0.9374$，$\hat{x}_{13,1}=-0.585$，$\hat{x}_{14,1}=0.6012$，

$\hat{x}_{15,1}=-1.2608$，$\hat{x}_{16,1}=-0.2082$。

因此

$$\hat{z}_{17}=\hat{x}_{16,1}+\bar{z}=-0.2082+9.95=9.7418 \text{（真值为9.6）},$$

$$\hat{x}_{17,1}=0.72(-0.35-(-0.2082))=-0.1021.$$

---

① 这一例题的数据选自[28]第82页的资料。

故

$$\hat{z}_{18}=\hat{x}_{17,1}+\bar{z}=9.848 \text{（真值为10.2）},$$

$$\hat{x}_{18,1}=0.72\times(0.25-(-0.1021))=0.2535.$$

因此

$$\hat{z}_{19}=\hat{x}_{18,1}+\bar{z}=10.20 \text{（真值为10.6）}.$$

(3) 关于 ARMA 序列的预报公式

设 x_t 是 ARMA(p,q)序列，其模型方程为

$$\sum_{l=0}^{p}\varphi_l x_{t-l}=\sum_{l=0}^{q}\theta_l\varepsilon_{t-l}\quad(\varphi_0=1)$$

或

$$\sum_{l=0}^{p}\varphi_l x_{t+\tau-l}=\sum_{l=0}^{q}\theta_l\varepsilon_{t+\tau-l}. \tag{3.83}$$

对(3.83)的两边求在 $H_x(t)$的投影，则得

$$\sum_{l=0}^{p}\varphi_l\hat{x}_{t,\tau-l}=\sum_{l=0}^{q}\theta_l E[\varepsilon_{t+\tau-l}|H_x(t)]. \tag{3.84}$$

当 $\tau>q$ 时

$$\varepsilon_{t+\tau-l}\perp H_x(t),\quad 0\leqslant l\leqslant q.$$

从而当 $\tau>q$ 时

$$\sum_{l=0}^{p}\varphi_l\hat{x}_{t,\tau-l}=0. \tag{3.85}$$

由此得

$$\hat{x}_{t,\tau}=-\sum_{l=1}^{p}\varphi_l\hat{x}_{t,\tau-l}. \tag{3.86}$$

表明当 $\tau>q$ 时，$\hat{x}_{t,\tau}$ 完全由$(\hat{x}_{t,\tau-1},\cdots,\hat{x}_{t,\tau-p})$诸预报值确定。

对于 $\tau\leqslant q$,

$$\hat{x}_{t+1,\tau}=c_\tau\varepsilon_{t+1}+\hat{x}_{t,\tau+1}, \tag{3.87}$$

$$\varepsilon_{t+1}=\frac{1}{c_0}(x_{t+1}-\hat{x}_{t,1}) \tag{3.88}$$

对 ARMA 模型同样成立，故有

$$\hat{x}_{t+1,\tau}=\hat{x}_{t,\tau+1}+\frac{c_\tau}{c_0}[x_{t+1}-\hat{x}_{t,1}],\quad 1\leqslant\tau<q. \tag{3.89}$$

$\tau<q$ 时预报公式和 MA 一样。但对 $\tau=q$ 时

$$\hat{x}_{t+1,\tau}=\hat{x}_{t,q+1}+\frac{c_q}{c_0}[x_{t+1}-\hat{x}_{t,1}]. \tag{3.90}$$

注意：(3.90)右端第一项在 MA 场合下 $=0$.现在利用(3.86)得

$$\hat{x}_{t,q+1}=-\sum_{l=1}^{p}\varphi_l\hat{x}_{t,q+1-l} \tag{3.91}$$

$$=\begin{cases}-\sum_{l=1}^{p}\varphi_l\hat{x}_{t,q+1-l}, & \text{当 } p\leqslant q;\\ -\sum_{l=1}^{q}\varphi_l\hat{x}_{t,q+1-l}+\sum_{l=q+1}^{p}\varphi_l x_{t+q+1-l}, & \text{当 } p>q.\end{cases} \tag{3.92}$$

(因为当 $p\geqslant l\geqslant q+1$ 时，$q+1-l\leqslant0$,则 $\hat{x}_{t,q+1-l}$ 等于 $x_{t+q+1-l}$.)
综合(3.89),(3.90),(3.92)则可得 ARMA 模型的向量预报公式：

定理3.6 设 x_t 为 ARMA(p,q)序列，则有以下的向量预报公式：

$$\hat{\boldsymbol{W}}_{t+1}^{q}=\boldsymbol{G}\hat{\boldsymbol{W}}_{t}^{q}+\boldsymbol{\alpha}x_{t+1}+\boldsymbol{\beta}, \tag{3.93}$$

其中

$$\boldsymbol{G}=\begin{bmatrix}-\frac{c_1}{c_0} & 1 & 0 & \cdots & 0\\ -\frac{c_2}{c_0} & 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \ddots & \vdots\\ -\frac{c_{q-1}}{c_0} & 0 & \cdots & & 1\\ -\frac{c_q}{c_0}-\varphi_q' & -\varphi_{q-1}' & \cdots & -\varphi_2' & -\varphi_1'\end{bmatrix}, \tag{3.94}$$

$$\varphi_j'=\begin{cases}\varphi_j, & 0<j\leqslant p,\\ 0, & j>p,\end{cases} \tag{3.95}$$

$$a=\left(\frac{c_1}{c_0},\frac{c_2}{c_0},\cdots,\frac{c_q}{c_0}\right)',\tag{3.96}$$

$$\beta=\left(0,0,\cdots,-\sum_{j=q+1}^{p}\varphi_j x_{t+q-j+1}\right)'.\tag{3.97}$$

（约定：当 $p\leqslant q$ 时 $\beta\equiv 0$。）

例3.6 设 x_t 为ARMA(1,2)模型，方程为

$$x_t-0.8x_{t-1}=\varepsilon_t-\varepsilon_{t-1}+0.24\varepsilon_{t-2}$$

求向量预报公式。

解 $p=1,q=2$，

$$\varphi_0=1,\varphi_1=-0.8;$$

$$\theta_0=1,\theta_1=-1,\theta_2=0.24,$$

可求得 Wold 系数为

$$c_0=1,\ c_1=\theta_1-\varphi_1c_0=-0.2,\ c_2=\theta_2-\varphi_1c_1=0.08,$$

$$\varphi_1'=\varphi_1=-0.8,\ \varphi_\tau'=0,\tau>1.$$

从而

$$\begin{bmatrix}\hat{x}_{t+1,1}\\ \hat{x}_{t+1,2}\end{bmatrix}=\begin{bmatrix}0.2 & 1\\ -0.08 & 0.8\end{bmatrix}\begin{bmatrix}\hat{x}_{t,1}\\ \hat{x}_{t,2}\end{bmatrix}+\begin{bmatrix}-0.2\\ 0.08\end{bmatrix}x_{t+1}.\tag{3.98}$$

（由于 $q>p$，则 $\beta\equiv 0$，无第三项。）

对于 $\tau>q=2$，用(3.86)则

$$\hat{x}_{t,\tau}=-\varphi_1\hat{x}_{t,\tau-1}=(-\varphi_1)^2x_{t,\tau-2}=\cdots$$

$$=(-\varphi_1)^{\tau-2}\hat{x}_{t,2}=(0.8)^{\tau-2}\hat{x}_{t,2}.$$

例3.7 设 x_t 为ARMA(2,1)模型，方程为

$$x_t+0.54x_{t-1}-0.3x_{t-2}=\varepsilon_t-0.5\varepsilon_{t-1},$$

求递推向量预报公式。

解 $p=2,q=1$，

$$\varphi_1=0.54,\varphi_2=-0.3,$$

$$\theta_0=1,\theta_1=-0.5,$$

Wold 系数

$$c_0=\theta_0=1,c_1=-1.04,c_2=0.86.$$

$$\hat{x}_{t+1,1}=\left(-\frac{c_1}{c_0}-\varphi_1\right)\hat{x}_{t,1}+\frac{c_1}{c_0}x_{t+1}-\varphi_2 x_t$$

$$=0.5\hat{x}_{t,1}-1.04x_{t+1}+0.3x_t.$$

当 $\tau\geqslant 2$ 时，由(3.86)

$$\hat{x}_{t+1,\tau}=-0.54x_{t+1,\tau-1}+0.3x_{t+1,\tau-2}.$$

§3 时间序列的线性滤波

1° 一般的讨论

设有二个平稳过程 x_t,y_t，具有性质

$$B_{xy}(t,\tau)=Ex_{t+\tau}y_t=B_{xy}(\tau),\qquad(3.99)$$

则称为 x_t 与 y_t 是平稳相关的。

设 x_t,y_t 平稳相关(零均值)，对 $\tau\geqslant 0$，称已知$\{x_n;\ n\leqslant t\}$之下对 $y_{t+\tau}$ 的预测性滤波乃指：寻求 $H_x(t)=\mathscr{L}\{x_n;\ n\leqslant t\}$中的一个元素 $\zeta_{t,\tau}$使得

$$\|y_{t+\tau}-\zeta_{t,\tau}\|^2=\inf_{\zeta\in H_x(t)}\|y_{t+\tau}-\zeta\|^2.\qquad(3.100)$$

由于 $y_{t+\tau}\in H_y=\mathscr{L}\{y_t;t=0,\pm1,\cdots\}\subset H$(见(3.10))，故 $\zeta_{t,\tau}$ 是 $y_{t+\tau}$ 对 $H_x(t)$的投影，它存在且唯一。由第一章的讨论知，$\zeta_{t,\tau}$ 可表成

$$\zeta_{t,\tau}=\int_{-\pi}^{\pi}e^{i\lambda t}H_{t,\tau}(\lambda)dZ_x(\lambda).\qquad(3.101)$$

重要的是要求出 $H_{t,\tau}(\lambda)$函数的显明表达式。所谓显明表达式是指将 $H_{t,\tau}(\lambda)$用谱函数 $F_{yy}(\lambda),F_{xy}(\lambda),F_{xx}(\lambda)$来表示，其中

$$F_{xy}(\lambda)=\frac{1}{2\pi}W_{xy}(\lambda+0)-c=\overline{F_{yx}(\lambda)}.\qquad(3.102)$$

而

$$W_{xy}(\lambda)=B_{xy}(0)\lambda-\sum_{k\neq 0}\frac{B_{xy}(k)}{\mathrm{i}k}\mathrm{e}^{-\mathrm{i}k\lambda}, \quad (3.103)$$

$$c=-\frac{1}{2\pi}W_{xy}(-\pi+0) \quad (3.104)$$

称为x_t,y_t的互谱函数①。$F_{yy}(\lambda),F_{xx}(\lambda)$为$y_t,x_t$的谱。(3.101)中的$H_{t,\tau}(\lambda)$我们称为$x_t$对$y_{t+\tau}$预测性滤波的谱特征，工程上称为滤波的频率响应函数，因为$H_{t,\tau}(\lambda)$给定之后，工程技术上可设计一个滤波器，使得将x_s $(s\leqslant t)$输入该滤波器后将产生最优的$\zeta_{t,\tau}$输出。

以下先讨论：假定已知全体观测资料$\{x_t;t=0,\pm 1,\cdots\}$，求对$y_{t+\tau}$的滤波问题。

定理3.7 设x_t,y_t平稳相关，互相关函数$R_{yx}(\tau)$满足

$$\sum_{k}|R_{yx}(\tau)|<+\infty, \quad (3.105)$$

则已知全体$\{x_t\}$元素下对$y_{t+\tau}$的最优滤波为

$$\zeta_{t,\tau}=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}\lambda t}\left(\mathrm{e}^{\mathrm{i}\lambda\tau}\frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}\right)\mathrm{d}Z_x(\lambda). \quad (3.106)$$

证明 由(3.105)及(3.103)，$f_{yx}(\lambda)$存在，

$$f_{yx}(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}R_{yx}(k)\mathrm{e}^{-\mathrm{i}k\lambda}. \quad (3.107)$$

$\zeta_{t,\tau}$显然是$y_{t+\tau}$在$H_x=\mathscr{L}\{x_t;t=0,\pm,\cdots\}$上的投影，它是存在唯一的。设

$$\zeta_{t,\tau}=\int_{-\pi}^{\pi}\varphi_{t,\tau}(\lambda)\mathrm{d}Z_x(\lambda), \quad (3.108)$$

则应有

$$(y_{t+\tau}-\zeta_{t,\tau},x_n)=0, \quad \forall n. \quad (3.109)$$

---

① 在第六章中，我们还要讨论多元随机序列的谱函数及其性质。如果$F_{xy}(\lambda)$有导函数$f_{xy}(\lambda)$，则称之为互谱密度函数。

由(3.109)可得

$$
\begin{aligned}
0 &= (y_{t+\tau}, x_n) - (\zeta_{t,\tau}, x_n) \\
&= \left(\int_{-\pi}^{\pi} e^{i(t+\tau)\lambda}\, dZ_y(\lambda), \int_{-\pi}^{\pi} e^{in\lambda} dZ_x(\lambda)\right) \\
&\quad - \left(\int_{-\pi}^{\pi} \varphi_{t,\tau}(\lambda) dZ_x(\lambda), \int_{-\pi}^{\pi} e^{in\lambda} dZ_x(\lambda)\right) \\
&= \int_{-\pi}^{\pi} e^{i(t+\tau-n)\lambda} f_{yx}(\lambda) d\lambda - \int_{-\pi}^{\pi} e^{-in\lambda} \varphi_{t,\tau}(\lambda) f_{xx}(\lambda) d\lambda
\end{aligned}
$$

对一切 n 成立，因此

$$
\varphi_{t,\tau}(\lambda) = e^{i(t+\tau)\lambda} \frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}. \quad | \tag{3.110}
$$

由(3.110)及(3.108)立刻可得到 τ 步预测性滤波的谱特征为

$$
H_\tau(\lambda) = e^{i\tau\lambda} \frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}. \tag{3.111}
$$

推论 设 $x_t = y_t + n_t$，y_t 与 n_t 独立，相关函数 $R_{yy}(\tau)$满足

$$
\sum_k |R_{yy}(k)| < +\infty, \tag{3.112}
$$

n_t 的谱密度为 $f_{nn}(\lambda)$，则已知$\{x_t, t=0, \pm 1, \cdots\}$之下，对 $y_{t+\tau}$的最优滤波为

$$
\zeta_{t,\tau} = \int_{-\pi}^{\pi} e^{i\lambda(t+\tau)} \frac{f_{yy}(\lambda)}{f_{yy}(\lambda) + f_{nn}(\lambda)} dZ_x(\lambda). \tag{3.113}
$$

证明 由于 n_t 与 y_t 独立，所以

$$
\begin{aligned}
Ex_{t+\tau}x_t &= Ey_{t+\tau}y_t + Ey_{t+\tau}n_t + En_{t+\tau}y_t + En_{t+\tau}n_t \\
&= Ey_{t+\tau}y_t + En_{t+\tau}n_t \text{①}
\end{aligned}
$$

---

① 因对平稳列我们皆假定均值为零。

$$= R_{yy}(\tau) + R_{nn}(\tau). \tag{3.114}$$

故

$$E y_{t+\tau} x_t = E y_{t+\tau} y_t + E y_{t+\tau} n_t = R_{yy}(\tau). \tag{3.115}$$

$$\begin{cases} f_{yx}(\lambda) = f_{yy}(\lambda), \\ f_{xx}(\lambda) = f_{yy}(\lambda) + f_{xx}(\lambda). \end{cases} \tag{3.116}$$

代入(3.110)即得。 ▌

以上结果虽然简单，但工程应用中却是非常重要的，它告诉我们已知全资料的噪声滤波器应该怎样设计。

假定 $x_t = y_t + n_t$ 是观测序列，y_t 是信号序列，n_t 是噪声序列，二者相互独立，设想 $f_{yy}(\lambda)$ 与 $f_{nn}(\lambda)$ 的谱密度如图 3.1 所示。噪声谱与信号谱在(α,β)区间上有交而$(0,\alpha)$和(β,π)不相交，则由(3.113)式知，谱特征(不妨考虑 $\tau=0$)$H_0(\lambda)$在$(0,\alpha)$纯噪声频段上应取为零；在(β,π)纯信号频段上应取为1，而在(α,β)上

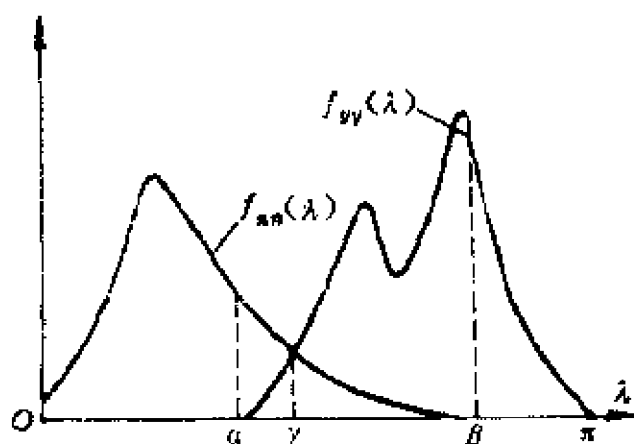


图 3.1

$$H_0(\lambda) = \frac{1}{1 + \frac{f_{nn}(\lambda)}{f_{yy}(\lambda)}} < 1.$$

故可得类似于图3.2的 $H_0(\lambda)$。

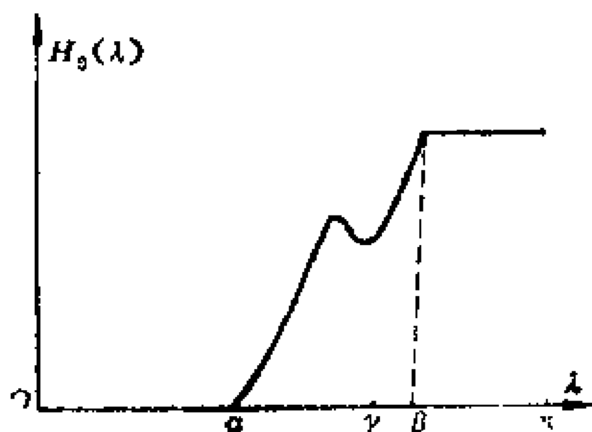


图 3.2

更进一步地，设想 $f_{nn}(\lambda)$ 与 $f_{yy}(\lambda)$ 完全无交，并且信号谱 $f_{yy}(\lambda)$ 集中在 $A\subset[-\pi,\pi]$，则

$$H_0(\lambda)=\begin{cases}1, & \lambda\in A\\ 0, & \lambda\overline{\in} A\end{cases} \tag{3.117}$$

$$=\chi_A(\lambda),\quad 0\leqslant\lambda\leqslant\pi.$$

故

$$\zeta_{t,0}=\int_{-\pi}^{\pi}e^{i\lambda t}H_0(\lambda)dZ_x(\lambda)=\int_{-\pi}^{\pi}e^{i\lambda t}\chi_A(\lambda)dZ_x(\lambda). \tag{3.118}$$

由 n_t 与 y_t 的正交性，不难得到

$$dZ_x(\lambda)=dZ_n(\lambda)+dZ_y(\lambda).$$

由于 $f_{nn}(\lambda)$ 与 $f_{yy}(\lambda)$ 无交，由一般$\|dZ(\lambda)\|^2=f(\lambda)d\lambda$知，$dZ_y(\lambda)$ 和 $dZ_n(\lambda)$ 有正测度的部分各自在 A 和 $\bar{A}$ 集上。故由(3.118)知

$$\zeta_{t,0}=\int_{-\pi}^{\pi}e^{i\lambda t}\chi_A(\lambda)dZ_y(\lambda)+\int_{-\pi}^{\pi}e^{i\lambda t}\chi_A(\lambda)dZ_n(\lambda)$$

$$=\int_{-\pi}^{\pi}e^{i\lambda t}dZ_y(\lambda)=y_t. \tag{3.119}$$

表明(3.113)的最优滤波完全从 x_t 中滤掉了噪声恢复了信号 y_t。

在应用中所谓已知全部观测 $\{x_s;s=0,\pm1,\cdots\}$ 求对 y_t 的滤波实际上是指对一段长观测

$$x_1,\ x_2,\ \cdots,\ x_{2N+1}\quad(N\text{充分大})$$

求对 y_N(中间)的滤波问题。运用(3.113)应事先对信号 y_t，n_t 分别进行谱分析——这些方法将在下一章中介绍。

以下讨论：已知半边数据$\{x_s; s\leqslant t\}$之下对$y_{t+\tau}$的最优滤波问题。这时可分两步投影，先用$y_{t+\tau}$对H_x投影得$\zeta_{t,\tau}$，然后再用$\zeta_{t,\tau}$对$H_x(t)$作投影得$\eta_{t,\tau}$。不难看出$\eta_{t,\tau}$就是$y_{t+\tau}$对$H_x(t)$的投影。分两步求解的原因是$\zeta_{t,\tau}$已属于H_x，它的谱特征由(3.110)为已知，而H_x中的标准正交基为Wold序列，于是容易求出对$H_t(\lambda)$的投影。

为了要求出$\eta_{t,\tau}$的滤波谱特征，需运用H_x与$L^2(\mathrm{d}F_{xx})$的同构对应。以下我们运用第二章中的方法。设x_t是ARMA序列，则有Wold列的表达式

$$\varepsilon_t=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}\lambda t}\Gamma_x^{-1}(\mathrm{e}^{-\mathrm{i}\lambda})\mathrm{d}Z_x(\lambda) \tag{3.120}$$

(见(2.51)式)。因此在$L^2(\mathrm{d}F_{xx})$中标准正交基为

$$\{\psi_t(\lambda)=\mathrm{e}^{\mathrm{i}\lambda t}\Gamma_x^{-1}(\mathrm{e}^{-\mathrm{i}\lambda})\,|\,t=0,\pm1,\cdots\}, \tag{3.121}$$

并且也具等时从属性质

$$\psi_t(\lambda)\in L_t^2(\mathrm{d}F_{xx})\ominus L_{t-1}^2(\mathrm{d}F_{xx}). \tag{3.122}$$

因此，$\zeta_{t,\tau}$依$\{\varepsilon_n\}$正交展开等价于它的谱特征$\mathrm{e}^{\mathrm{i}\lambda(t+\tau)}f_{yx}(\lambda)/f_{xx}(\lambda)$依$\{\psi_n(\lambda)\}$正交展开：

$$\mathrm{e}^{\mathrm{i}\lambda(t+\tau)}\frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}=\sum_{n=-\infty}^{\infty}\beta_{n,t}^{(\tau)}\psi_n(\lambda), \tag{3.123}$$

其中

$$\begin{aligned}\beta_{n,t}^{(\tau)}&=\left(\mathrm{e}^{\mathrm{i}\lambda(t+\tau)}\frac{f_{yx}(\lambda)}{f_{xx}(\lambda)},\mathrm{e}^{\mathrm{i}n\lambda}\Gamma_x^{-1}(\mathrm{e}^{-\mathrm{i}\lambda})\right)_{L^2(\mathrm{d}F_{xx})}\\&=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}\lambda(t+\tau-n)}\frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}\frac{1}{\overline{\Gamma_x(\mathrm{e}^{-\mathrm{i}\lambda})}}f_{xx}(\lambda)\mathrm{d}\lambda\\&=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}\lambda(t+\tau-n)}\frac{f_{yx}(\lambda)}{\overline{\Gamma_x(\mathrm{e}^{-\mathrm{i}\lambda})}}\mathrm{d}\lambda\\&=\beta_{t-n+\tau}.\end{aligned} \tag{3.124}$$

将(3.124)代入(3.123)可得(令$l=t-n$)

$$\mathrm{e}^{\mathrm{i}\lambda(t+\tau)}\frac{f_{yx}(\lambda)}{f_{xx}(\lambda)}=\sum_{l=-\infty}^{\infty}\beta_{l+\tau}\psi_{t-l}(\lambda). \tag{3.125}$$

因此对 $L_1^2(\mathrm{d}F_{xx})$的投影就是

$$\sum_{l=0}^{\infty}\beta_{l+\tau}\psi_{t-l}(\lambda). \tag{3.126}$$

由此知

$$\begin{aligned}\underset{H_x(t)}{\mathrm{Proj}}(\zeta_{t,\tau}) &= \int_{-\pi}^{\pi}\left(\sum_{l=0}^{\infty}\beta_{l+\tau}e^{i(t-l)\lambda}\,\Gamma_x^{-1}(e^{-i\lambda})\right)\mathrm{d}Z_x(\lambda)\\ &= \int_{-\pi}^{\pi}e^{it\lambda}\left(\sum_{l=0}^{\infty}\beta_{l+\tau}e^{-il\lambda}\right)\Gamma_x^{-1}(e^{-i\lambda})\mathrm{d}Z_x(\lambda).\end{aligned} \tag{3.127}$$

可见滤波谱特征为

$$H_\tau(\lambda)=\frac{\sum\limits_{l=0}^{\infty}\beta_{l+\tau}e^{-il\lambda}}{\Gamma_x(e^{-i\lambda})}, \tag{3.128}$$

其中 $\beta_{l+\tau}$ 由(3.124)确定。

综上所述，我们可得：

定理3.8 设 x_t 为 ARMA 序列，其方程为

$$\Phi(U)x_t=\Theta(U)\varepsilon_t.$$

又 y_t 是与 x_t 平稳相关的序列，互相关函数满足

$$\sum_k|R_{yx}(k)|<+\infty.$$

对 $\tau\geqslant0$，利用$\{x_k;k\leqslant t\}$对 $y_{t+\tau}$ 的最优线性滤波为

$$\eta_{t,\tau}=\int_{-\pi}^{\pi}e^{i\lambda t}H_\tau(\lambda)\mathrm{d}Z_x(\lambda), \tag{3.129}$$

其中

$$H_\tau(\lambda)=\left(\sum_{l=0}^{\infty}\beta_{l+\tau}e^{-il\lambda}\right)\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}, \tag{3.130}$$

$$\beta_{l+\tau}=\int_{-\pi}^{\pi}e^{i(l+\tau)\lambda}f_{yx}(\lambda)\left(\overline{\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}}\right)\mathrm{d}\lambda, \tag{3.131}$$

$$l=0,1,2,\cdots.$$

例3.8 设 x_t 和 y_t 平稳相关，它们的互谱密度和谱密度分别为

$$f_{yx}(\lambda) = C\frac{e^{-i\lambda}(1-\alpha e^{-i\lambda})}{e^{-i\lambda}-\beta},$$

$$f_{xx}(\lambda) = A^2\frac{|1-\alpha e^{-i\lambda}|^2}{|1-\beta e^{-i\lambda}|^2},$$

其中

$$0<\alpha,\beta<1,\quad A=\frac{1}{\sqrt{2\pi}}.$$

则

$$\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} = \Gamma_{xx}(e^{-i\lambda}) = \frac{1-\alpha e^{-i\lambda}}{1-\beta e^{-i\lambda}}.$$

故

$$f_{yx}(\lambda)\overline{\left(\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}\right)}$$

$$= C\frac{e^{-i\lambda}(1-\alpha e^{-i\lambda})}{e^{-i\lambda}-\beta}\left(\frac{1-\alpha e^{i\lambda}}{1-\beta e^{i\lambda}}\right)^{-1}$$

$$= C\frac{1-\alpha e^{-i\lambda}}{1-\alpha e^{i\lambda}}. \qquad (3.132)$$

然而

$$\frac{1-\alpha z^{-1}}{1-\alpha z} = \frac{1}{1-\alpha z} - \alpha\frac{1}{z}\frac{1}{1-\alpha z}$$

$$= \sum_{k=0}^{\infty}\alpha^k z^k - \alpha\sum_{k=0}^{\infty}\alpha^k z^{k-1}$$

$$= -\frac{\alpha}{z} + (1-\alpha^2)\sum_{k=0}^{\infty}\alpha^k z^k, \qquad (3.133)$$

$$0<|z|\leqslant 1,\quad 0<\alpha<1,$$

将 $z=e^{i\lambda}$ 代入(3.133)可得

$$f_{yx}(\lambda)\ \overline{\left(\frac{\Phi(e^{-i\lambda})}{\Theta(e^{-i\lambda})}\right)}$$

$$= C\left[-\alpha e^{-i\lambda} + (1-\alpha^2)\sum_{k=0}^{\infty}\alpha^k e^{ik\lambda}\right]. \qquad (3.134)$$

利用 Fourier 展式的唯一性，当 $\tau = 0$ 时(对应于(3.131))

$$\begin{cases} \beta_0 = B(1-\alpha^2); & B = 2\pi C, \\ \beta_1 = -B\alpha, & \\ \beta_k = 0, & k > 1. \end{cases}$$

可见

$$H_0(\lambda) = [B(1-\alpha^2) + (-B\alpha)e^{-i\lambda}]\frac{1-\beta e^{-i\lambda}}{1-\alpha e^{-i\lambda}}$$

$$= B[(1-\alpha^2) - \alpha e^{-i\lambda}][1-\beta e^{-i\lambda}]\frac{1}{1-\alpha e^{-i\lambda}}. \qquad (3.135)$$

然而令

$$\tilde{H}_0(z) = B[(1-\alpha^2) - \alpha z][1-\beta z]\frac{1}{1-\alpha z} \quad (|z| \leqslant 1).$$

则

$$\tilde{H}_0(z) = B\{(1-\alpha^2) - [(1-\alpha^2)\beta + \alpha]z + \alpha z\beta\}\sum_{k=0}^{\infty}\alpha^k z^k$$

$$= B\sum_{l=0}^{2}a_l z^l \sum_{k=0}^{\infty}\alpha^k z^k, \qquad (3.136)$$

其中

$$a_0 = 1-\alpha^2,\ \ a_1 = -(1-\alpha^2)\beta - \alpha,\ \ a_2 = \alpha\beta.$$

令 $\mu = k + l$，并交换求和次序，(3.136)可写为

$$\tilde{H}_0(z) = B\left\{(1-\alpha^2) - (\alpha^3 + \beta(1-\alpha^2))z + \sum_{\mu=2}^{\infty}(a_2\alpha^{\mu-2} + a_1\alpha^{\mu-1} + a_0\alpha^{\mu})z^{\mu}\right\}$$

$$= B\left\{ (1-\alpha^2) - (\alpha^3+\beta(1-\alpha^2))z + \alpha(\beta-\alpha)\sum_{\mu=2}^{\infty}\alpha^{\mu}z^{\mu}\right\}.$$

(3.137)

将上述 $\widetilde{H}_0(e^{-i\lambda}) = H_0(\lambda)$ 代回(3.129)则得 $\tau=0$ 之下的滤波公式

$$\eta_{t,0} = \int_{-\pi}^{\pi} e^{i\lambda t} H_0(\lambda) dZ_x(\lambda)$$

$$= B\left\{(1-\alpha^2)x_t - (\alpha^3+\beta(1-\alpha^2))x_{t-1} + \alpha(\beta-\alpha)\sum_{\mu=2}^{\infty}\alpha^{\mu}x_{t-\mu}\right\}.$$

(3.138)

对 τ 步预测性滤波，当 $\tau>1$，因 $\beta_{k+\tau}\equiv 0$，$k\geqslant 0$，故

$$\widetilde{H}_{\tau}(e^{-i\lambda}) \equiv 0. \qquad (3.139)$$

当 $\tau=1$，$\beta_1=-B\alpha$，$\beta_{k+1}=0$，$k\geqslant 1$，故

$$H_0(\lambda) = -B\alpha(1-\beta e^{-i\lambda})\frac{1}{1-\alpha e^{-i\lambda}}$$

$$= -\left(1+(\alpha-\beta)e^{-i\lambda}+\left(1-\frac{\beta}{\alpha}\right)\sum_{\mu=2}^{\infty}\alpha^{\mu}e^{-i\mu\lambda}\right)B\alpha.$$

$$\eta_{t,1} = -B\alpha\left\{x_t+(\alpha-\beta)x_{t-1}+\left(1-\frac{\beta}{\alpha}\right)\sum_{\mu=2}^{\infty}\alpha^{\mu}x_{t-\mu}\right\}. \qquad (3.140)$$

一般有理谱密度场合下 $x_t=y_t+n_t$，y_t 与 n_t 独立的噪声过滤(已知$\{x_s;s\leqslant t\}$)问题，亦可用类似于预测问题的 Яглом 方法来解(见[11])。

以上介绍的是滤波的频域方法，所用工具是 Wold 分解。以下介绍时域滤波的Kalman 方法，它的出发点是由状态方程和量测方程出发，优点是它的解是递推形式，适用于电子计算机来求解，并且有时不受平稳条件的限制。

2° Kalman 滤波

首先我们讨论 Kalman 滤波中的状态方程和量测方程的直观

意义。设想我们研究的对象是随机序列，它是AR(p)模型

$$x_k = -\sum_{l=1}^{p}\varphi_l x_{k-l} + w_k, \qquad (3.141)$$

其中$\{w_k\}$是独立同分布序列。采用第二章讨论的方法，可将(3.141)改写成

$$\begin{bmatrix} x_{k-p+1} \\ x_{k-p+2} \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & 0 \\ & & 0 & \ddots & \\ & 0 & & \ddots & 1 \\ -\varphi_p & -\varphi_{p-1} & \cdots & -\varphi_2 & -\varphi_1 \end{bmatrix}$$

$$\times \begin{bmatrix} x_{k-p} \\ x_{k-p+1} \\ \vdots \\ x_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} w_k \qquad (3.142)$$

或

$$\boldsymbol{X}_k = \boldsymbol{F}\boldsymbol{X}_{k-1} + \boldsymbol{\Gamma}\boldsymbol{w}_k, \qquad (3.143)$$

其中 $\boldsymbol{X}_k$ 是 p 维的列向量，$\boldsymbol{F}$ 是 $p\times p$ 的矩阵，$\boldsymbol{\Gamma}$ 是 $p\times r$ 的矩阵，$\boldsymbol{w}_k$ 是 r 维($r=1$)的列向量，(3.143)就称为模型方程，$\boldsymbol{w}_k$ 是模型噪声(白)。

设想我们研究的是$\{x_k\}$序列在空间中的传播，比方说散射通信，由于传播路径不同，达到接收点的时间不一，于是往往产生“多径效应”(见图3.3)。则在接收点(B)所收到的信号可表为

$$y_t = \sum_{k=0}^{p-1} h_k x_{t-k} + v_t, \qquad (3.144)$$

其中$\{v_t\}$为白噪声列，与 w_k 独立，$\{h_0, h_1, \cdots, h_{p-1}\}$为空间传播的

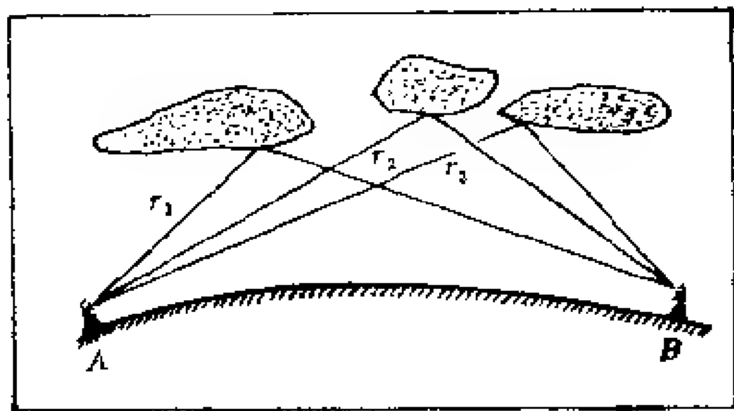


图 3.3

权系数。显然(3.144)可改写为

$$y_t=(h_{p-1},h_{p-2},\cdots,h_1,h_0)\begin{pmatrix}x_{t-p+1}\\ \vdots\\ x_{t-1}\\ x_t\end{pmatrix}+v_t$$

或

$$\boldsymbol{Y}_t=\boldsymbol{H}\boldsymbol{X}_t+\boldsymbol{v}_t, \tag{3.145}$$

其中 $\boldsymbol{Y}_t$ 是 m 维($m=1$)列向量，$\boldsymbol{H}$ 是 $m\times p$ 的矩阵，$\boldsymbol{v}_t$ 是 m 维($m=1$)的列向量。$\boldsymbol{Y}_t$ 的方程(3.145)称为量测方程，$\boldsymbol{v}_t$ 称为量测噪声。

以下我们讨论 Kalman 滤波的出发点就是一对方程

$$\begin{cases}\boldsymbol{X}_k=\boldsymbol{F}\boldsymbol{X}_{k-1}+\boldsymbol{\Gamma}\boldsymbol{w}_k,\\ \boldsymbol{Y}_k=\boldsymbol{H}\boldsymbol{X}_k+\boldsymbol{v}_k,\end{cases} \tag{3.146}$$

其中维数和矩阵的阶数的约定同上，此外关于 $\boldsymbol{w}_k,\boldsymbol{v}_s$ 有如下规定

$$\begin{cases}E\boldsymbol{w}_k=E\boldsymbol{v}_s=0,\\ E\boldsymbol{w}_k\boldsymbol{w}_j^*=\boldsymbol{Q}\delta_{k,j},\\ E\boldsymbol{v}_k\boldsymbol{v}_j^*=\boldsymbol{R}\delta_{k,j},\\ E\boldsymbol{w}_k\boldsymbol{v}_j^*=0,\end{cases} \tag{3.147}$$

并且已知初值 $\boldsymbol{X}_0$，$E\boldsymbol{X}_0=\mu_0$，$D(\boldsymbol{X}_0)=\boldsymbol{P}_0$(方差阵)，

$$E\boldsymbol{X}_0\boldsymbol{w}_k^*=E\boldsymbol{X}_0\boldsymbol{v}_k^*=0. \tag{3.148}$$

以下称满足上述约定的方程(3.146)为状态方程。

实用科技中关心的问题是：由 X_0, μ_0, P_0 和观测 $\{Y_j\}$ 出发，如何给出对 X_k 的估计？

假定第 $k-1$ 步对 X_{k-1} 的最优线性滤波估计已经得到，是 $\hat{X}_{k-1}$，又在第 k 个时刻观测到 Y_k，于是我们想用线性组合 $A\hat{X}_{k-1}+BY_k$ 来估计 X_k，而 A，B 的选择应使得方差阵

$$\Delta_k = E(X_k-(A\hat{X}_{k-1}+BY_k))(X_k-(A\hat{X}_{k-1}+BY_k))^*$$

达极小，所谓达极小是指对别的 $\bar{A}\hat{X}_{k-1}+\bar{B}Y_k$ 的误差阵 D，$D-\Delta_k$ 是非负定阵(即 $D\geqslant\Delta_k$)。

为了解决这一问题，注意以下事实：

(1) 可将 $X_k-A\hat{X}_{k-1}-BY_k$ 改写为

$$\begin{aligned}
X_k &- A\hat{X}_{k-1}-BY_k \\
&= FX_{k-1}+\Gamma w_k - A\hat{X}_{k-1} - B[H(FX_{k-1}+\Gamma w_k)+v_k] \\
&= (I-BH)[F(X_k-\hat{X}_{k-1})+\Gamma w_k]-Bv_k \\
&\quad +[(I-BH)F-A]\hat{X}_{k-1}.
\end{aligned} \tag{3.149}$$

(2) 记

$$P_k = E[(X_k-\hat{X}_k)(X_k-\hat{X}_k)^*], \tag{3.150}$$

$$\Pi_k = FP_{k-1}F^*+\Gamma Q\Gamma^*, \tag{3.151}$$

则

$$\begin{aligned}
\Delta_k &- (I-BH)\Pi_k(I-BH)^*+BRB^* \\
&+[(I-BH)F-A](E\hat{X}_{k-1}\hat{X}_{k-1}^*)[(I-BH)F-A]^*.
\end{aligned} \tag{3.152}$$

将(3.152)展开，再配方可得

$$\begin{aligned}
\Delta_k &= BH\Pi_kH^*B^*-BH\Pi_k-\Pi_kH^*B^*+\Pi_k+BRB^* \\
&\quad +[(I-BH)F-A](E\hat{X}_{k-1}\hat{X}_{k-1}^*)[(I-BH)F-A]^* \\
&= [B-\Pi_kH^*(H\Pi_kH^*+R)^{-1}](H\Pi_kH^*+R) \\
&\quad \times[B-\Pi_kH^*(H\Pi_kH^*+R)^{-1}]^* \\
&\quad +[\Pi_k-\Pi_kH^*(H\Pi_kH^*+R)^{-1}H\Pi_k]
\end{aligned}$$

$$+[(I-BH)F-A](E\hat{X}_{k-1}\hat{X}_{k-1}^*)[(I-BH)F-A]^* \quad (3.153)$$

$= \text{I} + \text{II} + \text{III}$。

(3) 在(3.153)的三部分中 I 属CDC^*型，其中

$$D=(H\Pi_k H^*+R)。\quad (3.154)$$

而 $R \geqslant 0$(见(3.147))，

$$\Pi_k = FP_{k-1}F^* + \Gamma Q\Gamma^*, \quad (3.155)$$

其中 $Q \geqslant 0$，故 $\Gamma Q\Gamma^* \geqslant 0$。又 $P_{k-1} \geqslant 0$($k-1$ 步的滤波误差阵)，$FP_{k-1}F^* \geqslant 0$，可见 $\Pi_k \geqslant 0$。由(3.154)$D \geqslant 0$，III亦是$\geqslant 0$的。

其次由(3.153)的II看，II与A，B无关，因此要Δ_k最小，只需让I，III两部分最小即可，而I，III皆$\geqslant 0$。

(4) 令(3.153)中的I，III $=0$，这只需选

$$B=\Pi_k H^*(H\Pi_k H^*+R)^{-1} \triangleq K_k, \quad (3.156)$$

$$A=(I-K_k H)F。\quad (3.157)$$

因此最优线性滤波为

$$\begin{aligned}\hat{X}_k &= (I-K_k H)F\hat{X}_{k-1} + K_k Y_k \\ &= F\hat{X}_{k-1} + K_k(Y_k - HF\hat{X}_{k-1})。\end{aligned} \quad (3.158)$$

从(3.158)结果看，这一滤波公式是很合理的，因为和模型方程(3.146)比较可知

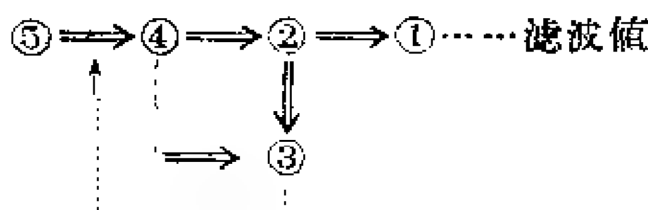
$$K_k(Y_k - HF\hat{X}_{k-1})$$

相当于误差噪声部分，而K_k正是调节增益系数。

综合以上结果，我们得到如下的针对状态方程(3.146)的Kalman滤波公式：

$$\left.\begin{aligned}
&\hat{X}_k = F\hat{X}_{k-1} + K_k(Y_k - HF\hat{X}_{k-1}), && ①\\
&K_k = \Pi_k H^*(H\Pi_k H^*+R)^{-1}, && ②\\
&P_k = (I-K_k H)\Pi_k, && ③\\
&\Pi_k = FP_{k-1}F^* + \Gamma Q\Gamma^*, && ④\\
&P_0 = D(X_0)。 && ⑤
\end{aligned}\right\} \quad (3.159)$$

而递推算法的流程是



一开始 P_0 用 $D(X_0)$ 而 $\hat{X}_0$ 用 0 向量来代替。设想 R,Q 是一维的，则由②—④看

(1) 当 R 值大时，K_k 变小，表明在量测方程中，若量测噪声大，滤波调节增益要小一些。

(2) 当 Q 值大时，Π_k 增大，由②式看 K_k 通常也会加大。

在上述流程中，出发点是 P_0，而 P_0 一般在实际工作中预先往往不知道，因而是由实测资料估计出来的。于是一个重要的问题是，如果 P_0 估计有些偏差，会不会愈递推误差愈大？以下定理告诉我们在一定条件下是可以避免的。

定理3.9 设状态方程(3.146)中的 F,H,Γ 满足以下不等式

$$\sum_{k=0}^{n-1} F^k\Gamma\Gamma^*(F^k)^*>0,$$

$$\sum_{k=0}^{n-1} (F^k)^*H^*HF^k>0, \tag{3.160}$$

则存在一个唯一的正定阵 P 使得对任意的初始方差 P_0，恒有 $P_k\to P$，$k\to\infty$，其中 n 为状态的维数。

这一定理的证明可参见于[27]。这一定理的重要性在于它告诉我们：在适当条件下，不论如何选初值，当递推次数足够长之后，它的滤波误差阵将趋于一个正定阵 P，从而它的增益矩阵也将趋于一个确定的增益阵，实用中表明递推到一定次数后即可免去对 K_k 的递推计算，因而大大减少了滤波的计算量。当然，许多应用工作者也常有一开始就取 K_k 为适当的一个常数阵，效果也不错。

最后要指出：上述滤波方法亦可用于状态方程为

$$\begin{cases} X_k = F_k X_{k-1} + \Gamma_k w_k, \\ Y_k = H_k X_k + v_k \end{cases} \tag{3.161}$$

的非定常系数的场合，这时数学上相当于 X_k 是非平稳列的滤波问题。

§4 与预测和滤波有关的问题

1° 一般平稳序列的预测与滤波

设 x_t 是（零均值）平稳列，它有谱密度 $f_{xx}(\lambda)$ 满足正则性条件

$$\int_{-\pi}^{\pi} \ln f_{xx}(\lambda) \mathrm{d}\lambda > -\infty \tag{3.162}$$

(ARMA 模型显然满足正则性条件)，我们来讨论这种比较一般条件下已知$\{x_s;\ s \leqslant t\}$对 $x_{t+\tau}$ 的预测问题。不失一般性，可利用平稳性将问题提为已知$\{x_s;\ s \leqslant 0\}$对 x_τ 的预测问题。

利用第二章的方法，可以证明在正则性条件下 x_t 必有 Wold 分解(见[13])

$$x_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}, \tag{3.163}$$

并且

$$c_k = (x_t, \varepsilon_{t-k}), \quad k \geqslant 0. \tag{3.164}$$

Wold列$\{\varepsilon_t\}$由(2.72)定义并具定理2.5 的 A,B,C,D 的性质。显见，已知$\{x_s;\ s \leqslant 0\}$对 x_τ 的预测量是

$$\underset{H_x(0)}{\mathrm{Proj}}(x_\tau) = \sum_{k=\tau}^{\infty} c_k \varepsilon_{\tau-k}, \tag{3.165}$$

并且

$$\varepsilon_t = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda t} \Gamma_{xx}^{-1}(\mathrm{e}^{-\mathrm{i}\lambda}) \mathrm{d}Z_x(\lambda), \tag{3.166}$$

$$\Gamma_{xx}(z)=\sum_{k=0}^{\infty} c_k z^k,\quad |z|<1, \tag{3.167}$$

其中 $\Gamma_{xx}(z)$ 是单位圆内的解析函数，满足

$$(1)\quad \sum_{k=0}^{\infty} |c_k|^2<+\infty; \tag{3.168}$$

$$(2)\quad \frac{1}{2\pi}|\Gamma_{xx}(e^{-i\lambda})|^2=f_{xx}(\lambda),\qquad -\pi\leqslant\lambda\leqslant\pi; \tag{3.169}$$

(3) $\Gamma_{xx}(z)$ 在单位圆内无零点。

这些公式和结论在第二章的 ARMA 模型中已经获得过因而是不陌生的。关键的问题是在给定 x_t 的谱密度 $f_{xx}(\lambda)$ 之后，如何求Wold 系数$\{c_k\}$和 $\Gamma_{xx}(z)$ 这个极关重要的极大解析函数。Колмогоров 在[13]中证明了：对正则的 x_t 序列，

$$\Gamma_{xx}(z)=\sqrt{2\pi}\ e^{\frac{1}{4\pi}\int_{-\pi}^{+\pi}\ln f_{xx}(\lambda)\frac{e^{-i\lambda}+z}{e^{-i\lambda}-z}d\lambda} \tag{3.170}$$

$$=\sqrt{2\pi}\ e^{\left(\frac{a_0^{(x)}}{2}+\sum_{k=1}^{\infty}(a_k^{(x)}+i\,b_k^{(x)})\,z^k\right)}, \tag{3.171}$$

其中 $|z|<1$，$\{a_k^{(x)},b_k^{(x)}\}$ 为 $\frac{1}{2}\ln f_{xx}(\lambda)$ 的 Fourier 系数，即

$$\frac{1}{2}\ln f_{xx}(\lambda)\sim\frac{a_0^{(x)}}{2}+\sum_{k=1}^{\infty}(a_k^{(x)}\sin kx+b_k^{(x)}\cos kx). \tag{3.172}$$

当 x_t 满足

$$\Phi(U)x_t=\Theta(U)\varepsilon_t$$

这种 ARMA 模型的特殊条件下，不难证明(见[29])，上述极大解析函数正是：

$$\Gamma_{xx}(z)=\frac{\Theta(z)}{\Phi(z)},\quad |z|<1.$$

在应用中要求(3.171)的 Taylor 展开系数$\{c_k\}$ 当然还是很不

容易的，如果是 ARMA 模型自然可以用第二章中的递推公式(2.81)逐步推出$\{c_k\}$。

利用(3.167)不难看出

$$x_t=\int_{-\pi}^{\pi}\left(\sum_{k=0}^{\infty}c_k\,\mathrm{e}^{-ik\lambda}\right)\mathrm{e}^{i\lambda t}\mathrm{d}Z_\varepsilon(\lambda)$$

$$=\int_{-\pi}^{\pi}\mathrm{e}^{i\lambda t}\Gamma_{xx}(\mathrm{e}^{-i\lambda})\mathrm{d}Z_\varepsilon(\lambda). \qquad (3.173)$$

由 (3.164)

$$c_k=\left(\int_{-\pi}^{\pi}\mathrm{e}^{i\lambda t}\Gamma_{xx}(\mathrm{e}^{-i\lambda})\mathrm{d}Z_\varepsilon(\lambda),\int_{-\pi}^{\pi}\mathrm{e}^{i\lambda(t-k)}\mathrm{d}\,Z_\varepsilon(\lambda)\right)$$

$$=\int_{-\pi}^{\pi}\mathrm{e}^{ik\lambda}\Gamma_{xx}(\mathrm{e}^{-i\lambda})f_{\varepsilon\varepsilon}(\lambda)\mathrm{d}\lambda$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{ik\lambda}\Gamma_{xx}(\mathrm{e}^{-i\lambda})\mathrm{d}\lambda,\quad k=0,1,2,\cdots. \qquad (3.174)$$

因此在获得$\Gamma_{xx}(\mathrm{e}^{-i\lambda})$后亦可用(3.174)Fourier 系数公式来求(比如用 FFT 等近代工具)。

以下我们容易得到关于一步预报误差的表达式。

定理3.10 设x_t是正则平稳序列，则

$$\|x_{t+1}-\hat{x}_{t+1}\|=c_0=\sqrt{2\pi}\;\mathrm{e}^{\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)\mathrm{d}\lambda}>0, \qquad (3.175)$$

其中

$$\hat{x}_{t+1}=E[x_{t+1}|H_x(t)]$$

为已知$\{x_s;\ s\leqslant t\}$对x_{t+1}的预报。

证明 由 Wold 分解

$$\hat{x}_{t+1}=E[x_{t+1}|H_x(t)]=\sum_{k=1}^{\infty}c_k\varepsilon_{t+1-k}. \qquad (3.176)$$

因而

$$\|x_{t+1}-\hat{x}_{t+1}\|^2=c_0^2.$$

然而由(3.170)看

$$c_0=\Gamma_{xx}(0)=\sqrt{2\pi}\ e^{\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)d\lambda}, \quad (3.177)$$

由正则性

$$\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)d\lambda>-\infty,$$

故 $c_0>0$。 ▎

在 ARMA 模型条件下，第二章中已得到

$$c_0=\theta_0=\sqrt{2\pi}\ e^{\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)d\lambda}, \quad (3.178)$$

因此

$$\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)d\lambda=2\pi\ln\frac{(\theta_0^2)}{2\pi}. \quad (3.179)$$

(3.179)告诉我们一个重要事实，即：$\ln f_{xx}(\lambda)$ 的积分值与一步预测误差的对数 $\ln c_0^2$ 是成正比关系的。

至于滤波问题也很容易用上述 Колмогоров 方法来解决。设 x_t, y_t 平稳相关（零均值）有谱密度 $f_{xx}(\lambda)$和互谱密度$f_{yx}(\lambda)$。已知$\{x_s;\ s\leqslant t\}$要对$y_{t+\tau}$进行预测性滤波，利用平稳性，只要求出已知$\{x_s;\ s\leqslant 0\}$对 y_τ 的滤波谱特征即可。由前知，我们要求 x_t 是正则序列，即

$$\int_{-\pi}^{\pi}\ln f_{xx}(\lambda)d\lambda>-\infty,$$

因而(3.119)至(3.128)的结论都是对的，因为在正则序列的条件下第二章中 $H_x(t)$和 $L_t^2(d\Gamma_{xx})$的同构对应及有关 Wold 序列的结论（只要不涉及 ARMA 模型参数）都成立，因而滤波谱特征是

$$H_\tau(\lambda)=\sum_{i=0}^{\infty}\beta_i^{(\tau)}e^{-i i\lambda}/\Gamma_x(e^{-i\lambda}),$$

而

$$\beta_{\tau+l}=\int_{-\pi}^{\pi} e^{i(l+\tau)\lambda}\left(\Gamma_{xx}^{-1}(e^{-i\lambda})\right)f_{yx}(\lambda)d\lambda,$$

$$l=0,1,2,\cdots,$$

由(3.170)知 $\Gamma_{xx}(e^{-i\lambda})$ 是

$$\Gamma_{xx}(z)=\sqrt{2\pi}\ e^{\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln f_{xx}(\mu)\frac{e^{-i\mu}+z}{e^{-i\mu}-z}d\mu},\qquad |z|<1$$

在 $z\to e^{-i\lambda}$ 的边值。

2° 关于非平稳列的平稳化处理及其预测

本书所讨论的问题绝大部分都以平稳列为对象，但是在应用中大量问题都涉及到非平稳过程和非平稳序列，以下只介绍在许多场合下要用到的简单的平稳化处理手段，这些方法虽然经常很有效但并不是一切非平稳序列经过以下处理后就一定能够平稳化。

假定 x_t 的记录如图3.4和3.6，则可认为它们分别有线性和二次曲线的趋势变化。然而由一次和二次曲线的性质易知经过一阶和二阶差分处理之后它们都变成了常数(见图3.5和3.7)。

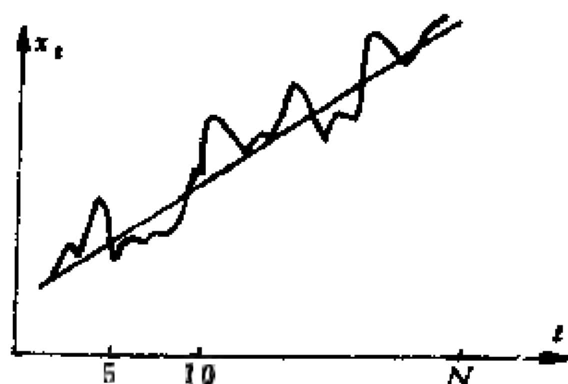


图 3.4

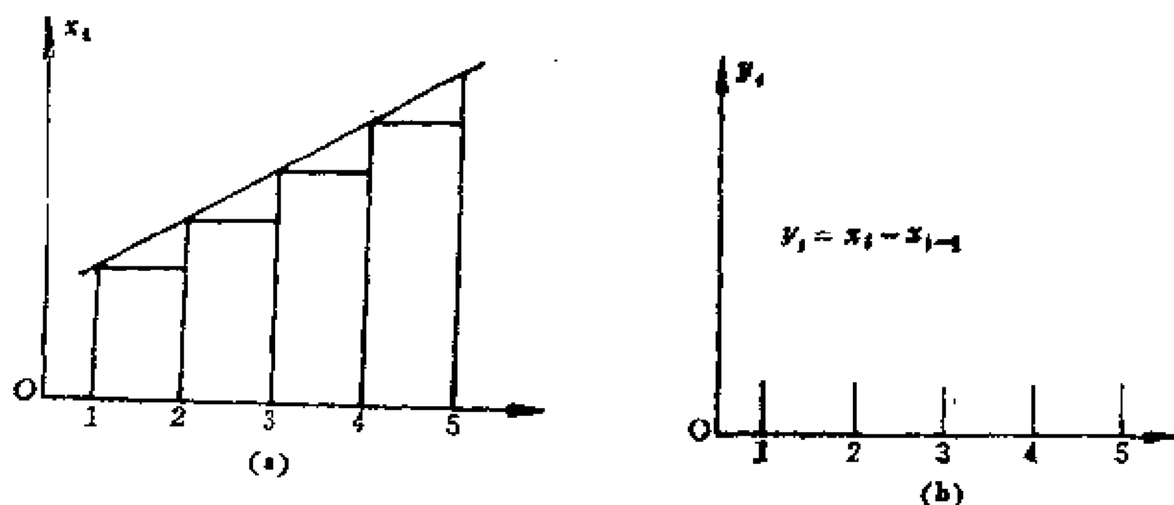


图 3.5

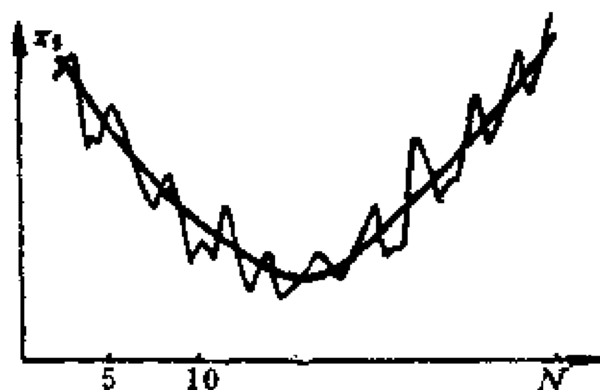


图 3.6

一般说来，许多非平稳列 Z_t，经过 d 阶差分后可得一平稳列 x_t:

$$x_t \triangleq \nabla^d Z_t = \sum_{k=0}^{d} (-1)^k C_d^k Z_{t-k}. \qquad (3.180)$$

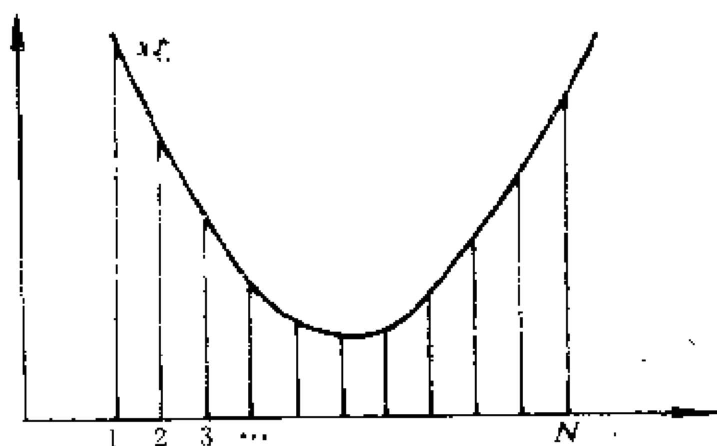
如果它还是平稳ARMA(p,q)模型，我们就称Z_t是ARIMA(p,d,q)模型。确切的定义如下(见[26])：

设 Z_t 是随机列，它满足以下条件

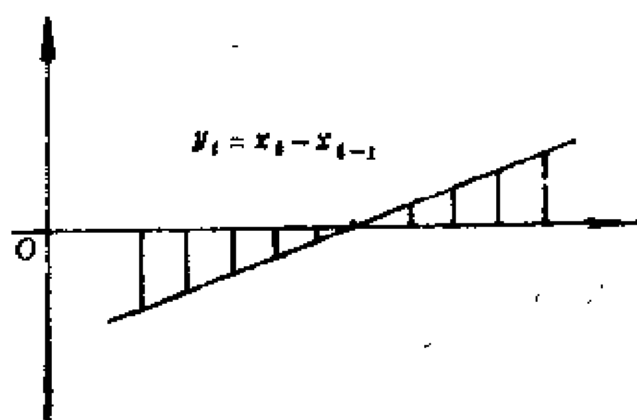
(1) $E|Z_t|^2 < +\infty$;

(2) 存在正整数 d 使得

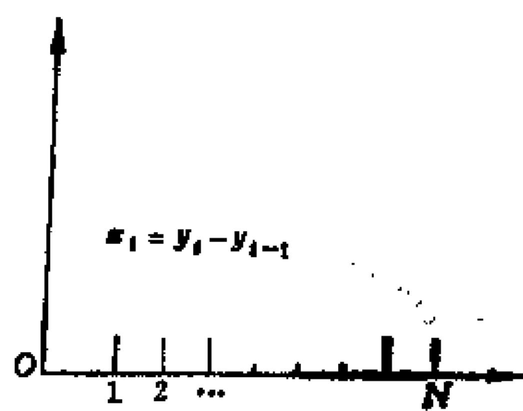
$$EZ_t = 0, \quad t = 1,2,3,\cdots,d; \qquad (3.181)$$



(a)



(b)



(c)

图 3.7

并且

$$x_t = \sum_{k=0}^{d}(-1)^k C_d^k Z_{t-k} \tag{3.182}$$

满足 ARMA(p,q)模型；

(3) $EZ_s x_t = 0,\quad 1 \leqslant s \leqslant d, d < t$。 (3.183)

则称 Z_t 为 ARIMA(p,d,q)模型。

以上是 x_t 通过$\nabla^d Z_t$（即 d 阶差商（3.182)式）来表示；反之，Z_t 亦可通过 x_t （$t>d$)及 d 个 Z_t 的初值 $Z_1, Z_2, \cdots, Z_d$ 来表示。

例如：对$d=1$，则

$$Z_t = (Z_t - Z_{t-1}) + (Z_{t-1} - Z_{t-2}) + \cdots + (Z_2 - Z_1) + Z_1 \tag{3.184}$$

$$= Z_1 + \sum_{j=1}^{t-1} x_{1+j},\quad t > d = 1。 \tag{3.185}$$

当然也可以用

$$Z_t = Z_k + \sum_{l=1}^{t-k} x_{k+l},\quad t > k \tag{3.186}$$

来表示。

对$d=2$，运用(3.186)形式($t>k$)

$$\begin{aligned}
Z_t &= Z_k + \sum_{l=1}^{t-k} y_{k+l} \qquad (y_{k+l} = Z_{k+l} - Z_{k+l-1}) \\
&= Z_k + \sum_{l=1}^{t-k}\left(y_k + \sum_{j=1}^{l} x_{k+j}\right) \\
&= Z_k + (t-k)y_k + \sum_{l=1}^{t-k}\sum_{j=1}^{l} x_{k+j} \\
&= Z_k + (t-k)(Z_k - Z_{k-1}) + \sum_{l=1}^{t-k}\sum_{j=1}^{l} x_{k+j} \\
&= Z_k + (t-k)(Z_k - Z_{k-1}) + \sum_{j=1}^{t-k}(t-k-j+1)x_{k+j}。
\end{aligned} \tag{3.187}$$

特别可取 $k=2$，则对 $t>2$

$$Z_t=Z_2+(t-2)(Z_2-Z_1)+\sum_{j=1}^{t-2}(t-j-1)x_{2+j}$$

$$=Z_2+C_{t-2}^1\nabla Z_2+\sum_{j=1}^{t-2}C_{t-j-1}^1 x_{2+j}. \tag{3.188}$$

为方便起见，Z_t对 d 阶差分，往往取 $Z_1,Z_2,\cdots,Z_d$以及 x_t 来表示：

$$Z_t=\sum_{i=0}^{d-1}C_{t-k+i-1}^i\nabla^i Z_k+\sum_{j=1}^{t-k}C_{t-k-j+d-1}^{d-1}x_{k+j}\quad(t>k\geqslant d) \tag{3.189}$$

当$k=d$时，

$$Z_t=\sum_{i=0}^{d-1}C_{t-d+i-1}^i\nabla^i Z_d+\sum_{j=1}^{t-d}C_{t-j-1}^{d-1}x_{d+j}. \tag{3.190}$$

由于 x_t 是 ARMA(p,q)模型，故 Z_t 的预报可通过(3.190) 的关系式给出。

定理3.11 设 Z_t 为 ARIMA(p,d,q)序列，则已知 $\{Z_k|k\leqslant t\}$ 之下关于 $Z_{t+\tau}$ 的最优预报$(t>d,\tau>0)$

$$E[Z_{t+\tau}|Z_1,Z_2,\cdots,Z_d;\ x_t,x_{t-1},\cdots]$$
$$=E[Z_{t+\tau}|\mathscr{Z}_t]$$
$$=\sum_{i=0}^{d-1}C_{\tau+i-1}^i\nabla^i Z_t+\sum_{j=1}^{\tau}C_{\tau-j+d-1}^{d-1}\hat{x}_{t,j}, \tag{3.191}$$

其中

$$\hat{x}_{t,j}=E[x_{t+j}|H_x(t)],$$

而(3.191)左端是指 $Z_{t+\tau}$ 对

$$\mathscr{Z}_t=\left\{\zeta:\ \zeta=w+\sum_{j=1}^{d}C_jZ_j,\ w\in H_x(t),C_j任意\right\} \tag{3.192}$$

的投影。

证明 利用(3.189)，$t+\tau>k\geqslant d$，再取$k=t$，得

$$\begin{aligned} Z_{t+\tau} &= \sum_{i=0}^{d-1} C^{i}_{t+\tau-k+i-1}\nabla^{i} Z_k + \sum_{j=1}^{t+\tau-k} C^{d-1}_{t+\tau-j-k+d-1} x_{k+j} \\ &= \sum_{i=0}^{d-1} C^{i}_{\tau+i-1}\nabla^{i} Z_t + \sum_{j=1}^{\tau} C^{d-1}_{\tau-j+d-1} x_{t+j}. \end{aligned} \tag{3.193}$$

对(3.193)在$\mathscr{Z}_t$上投影，得

$$\begin{aligned} E[Z_{t+\tau}\mid\mathscr{Z}_t] &= \sum_{i=0}^{d-1} C^{i}_{\tau+i-1} E[\nabla^{i} Z_t\mid\mathscr{Z}_t] \\ &\quad + \sum_{j=1}^{\tau} C^{d-1}_{\tau-j+d-1} E[x_{t+j}\mid\mathscr{Z}_t]. \end{aligned} \tag{3.194}$$

由于

$$\begin{gathered} \nabla^{i} Z_t = (1-U)^{i} Z_t, \qquad i=0,1,\cdots d-1, \\ \nabla^{i} Z_t \in \mathscr{Z}_t \ \text{(见(3.192)定义)}. \end{gathered} \tag{3.195}$$

又由约定 $x_t \perp Z_1, Z_2, \cdots, Z_d$，$t>d$ 所以

$$E[x_{t+j}\mid\mathscr{Z}_t] = E[x_{t+j}\mid H_x(t)] = \hat{x}_{t,j}.$$

代回(3.194)得

$$\begin{aligned} E[Z_{t+\tau}\mid\mathscr{Z}_t] &= \sum_{i=0}^{d-1} C^{i}_{\tau+i-1}\nabla^{i} Z_t \\ &\quad + \sum_{j=1}^{\tau} C^{d-1}_{\tau-j+d-1}\hat{x}_{t,j}. \quad \blacksquare \end{aligned} \tag{3.196}$$

除了上述ARIMA(p,d,q)以外，还可以考虑趋势变化有明显周期性的随机序列，如某城市的月平均温度，地球自转速度的变化，地区性的用电量等都有明显的季节性(周期)的变化。

季节性模型：称 x_t 为具有周期 s，阶数为$(p,d,q)\times(P,D,Q)$的季节性模型，假若它满足以下的方程

$$\varphi(U)\Phi(U^{s})\nabla^{d}\nabla_{s}^{D} x_t = \theta(U)\Theta(U^{s})\varepsilon_t, \tag{3.197}$$

其中 s 为正整数，

$$\nabla_s \triangleq 1-U^s, \quad \nabla_s^D \triangleq (1-U^s)^D,$$

$$\nabla^d \triangleq (1-U)^d, \tag{3.198}$$

$$\Phi(U^s) \triangleq 1+\Phi_1 U^s+\Phi_2 U^{2s}+\cdots+\Phi_P U^{Ps}, \tag{3.199}$$

$$\Theta(U^s) \triangleq \Theta_0+\Theta_1 U^s+\Theta_2 U^{2s}+\cdots+\Theta_Q U^{Qs}, \tag{3.200}$$

$$\varphi(U) \triangleq 1+\varphi_1 U+\varphi_2 U^2+\cdots+\varphi_p U^p, \tag{3.201}$$

$$\theta(U) \triangleq \theta_0+\theta_1 U+\theta_2 U^2+\cdots+\theta_q U^q, \tag{3.202}$$

且 $\Phi(Z^s)$, $\varphi(Z)$, $\Theta(Z^s)$, $\theta(Z)$ 的根皆在单位圆外，而 ε_t 是白噪声。

例3.9 x_t 是 $(p,0,q)\times(0,1,0)_s$（周期为 s）的模型，则方程(3.197)对应于

$$\varphi(U)\nabla_s x_t=\Theta_0\theta(U)\varepsilon_t, \tag{3.203}$$

即

$$y_t=\nabla_s x_t=(1-U^s)x_t=x_t-x_{t-s} \tag{3.204}$$

是 ARMA(p,q) 模型。

例3.10 x_t 是 $(p,1,q)\times(0,1,0)_s$ 模型，则(3.197)方程对应于

$$\varphi(U)\nabla\nabla_s x_t=\theta(U)\varepsilon_t, \tag{3.205}$$

即

$$\begin{aligned} y_t=\nabla\nabla_s x_t &= (1-U)(1-U^s)x_t \\ &= (1-U-U^s+U^{s+1})x_t \\ &= x_t-x_{t-1}-x_{t-s}+x_{t-s-1} \\ &= (x_t-x_{t-s})-(x_{t-1}-x_{t-s-1}) \\ &= Z_t-Z_{t-1} \end{aligned} \tag{3.206}$$

是 ARMA(p,q) 模型。(3.206)表明对原始数据 $\{x_t\}$ 先按周期 s（间隔 s）作差分得 Z_t，而后 Z_t 的一阶差分为 y_t，它是 ARMA(p,q)，也可说 Z_t 是 ARIMA$(p,1,q)$。

由上可见，若在(3.197)中，令

$$Z_t=\nabla^d\nabla_s^D x_t \tag{3.207}$$

并合并

$$\varphi\Phi = \bar{\Phi}, \quad \theta\Theta = \bar{\Theta} \tag{3.208}$$

两个多项式，则$\bar{\Phi}$的阶为$p+Ps$，$\bar{\Theta}$的阶为$q+Qs$，因而Z_t是ARMA$(p+Ps, q+Qs)$序列，但却是稀疏系数的。

类似于ARIMA模型的方法，在得到Z_t的预报后不难还原到x_t的预报公式，只是公式比起ARIMA模型更繁就是了(可参看[26])。

然而，如果x_t的周期性趋势很明显或明确有固定的(非随机型)的趋势项且难以用一个函数确切表达时，这类非平稳序列的平稳化应该用第五章我们将介绍的潜周期检测法来确定趋势项，其方法如下：

假设模型为

$$x_t = \sum_{k=0}^{p} c_k e^{-it\lambda_k} + \xi_t \tag{3.209}$$

或

$$x_t = \sum_{k=0}^{p} a_k \sin(t\lambda_k + \theta_k) + \xi_t, \tag{3.210}$$

其中$p, a_k, \lambda_k, \theta_k$皆未知，从$x_t$中可以估计出它们来，于是可令

$$\Delta_t = x_t - \sum_{k=0}^{\hat{p}} \hat{a}_k \sin(t\hat{\lambda}_k + \hat{\theta}_k), \tag{3.211}$$

而认为Δ_t是平稳过程，于是可以给它一个适当的ARMA模型的拟合。设为ξ_t，于是$x_{t+\tau}$的预报公式为

$$\hat{x}_{t+\tau} = \sum_{k=0}^{\hat{p}} \hat{a}_k \sin((t+\tau)\hat{\lambda}_k + \hat{\theta}_k) + \hat{\xi}_{t,\tau}. \tag{3.212}$$

而$\hat{\xi}_{t,\tau}$预报公式在前两节中已介绍过，特别用第四章中的极大熵方法很容易得到ξ_t的模型方程和预报$\hat{\xi}_{t,\tau}$。

3° 关于非均方准则下的最优滤波问题

在以前各节中我们介绍的问题其求解都是在均方准则(即

Hilbert 空间H的范数)意义下得到的，但是在许多工程技术问题中用均方准则并不适宜(见[30])而需要用别的准则——比如，一种比较重要而且在电讯工程界中已用了多年的是"极大信噪比"准则(见[17])，例如在雷达信号检测中，有时并不关心完全信号的恢复而首先是在强噪声背景中发现有回波信号，因而滤波器应尽可能地突出信号的存在。

为了从数学上能更本质地理解极大信噪比滤波以下先做一些准备工作。

设收到的讯号为

$$\zeta(t)=\xi(t)+s(t), \tag{3.213}$$

其中$\xi(t)$是均值为零的宽平稳列，$s(t)$是确定型的信号，可表为

$$s(t)=\int_{-\pi}^{\pi} e^{i\lambda t}\, dP(\lambda), \tag{3.214}$$

$P(\lambda)$是$[-\pi,\pi]$上的有界变差函数，引用以前记号，记

$$\begin{aligned} H_\zeta(\tau)&=\mathscr{L}\{\zeta(t)\mid t\leqslant\tau\},\\ H_\zeta&=\mathscr{L}\{\zeta(t)\mid t=0,\pm1,\cdots\}, \end{aligned} \tag{3.215}$$

于是有以下引理：

引理3.1 设$x\in H_\zeta(\tau)$，且$x=\lim\limits_{n\to\infty}(\xi_x^{(n)}+s_x^{(n)})$，其中

$$\begin{aligned} \xi_x^{(n)}&=\sum_{j=1}^{k_n}a_j^{(n)}\xi(t_j^{(n)}),\\ s_x^{(n)}&=\sum_{j=1}^{k_n}a_j^{(n)}s(t_j^{(n)}), \qquad t_j^{(n)}\leqslant\tau. \end{aligned} \tag{3.216}$$

则极限

$$s_x=\lim_{n\to\infty}s_x^{(n)},\quad \xi_x=\lim_{n\to\infty}\xi_x^{(n)} \tag{3.217}$$

分别存在，且有

$$x=\xi_x+s_x. \tag{3.218}$$

反之，若有$\{a_j^{(n)}\}$和$\{t_j^{(n)}|t_j^{(n)}\leqslant\tau\}$使(3.217)成立，则(3.218)的$x\in H_\xi(\tau)$。

证明 因为

$$\begin{aligned}&\|(\xi_x^{(n)}+s_x^{(n)})-(\xi_x^{(m)}+s_x^{(m)})\|^2\\&=\|\xi_x^{(n)}-\xi_x^{(m)}\|^2+\|s_x^{(n)}-s_x^{(m)}\|^2,\end{aligned} \tag{3.219}$$

故引理的结论显明成立。 ▎

由(3.218)分解的唯一性

$$(S/N)_\tau=\sup_{x\in H_\xi(\tau)}\frac{|s_x|}{\sqrt{E|\xi_x|^2}}, \tag{3.220}$$

称为τ时刻的信噪比。

由$H_\xi(\tau)\subset H_\xi(\tau+1)$知$(S/N)_\tau$是$\tau$的非降函数。我们关心的是：噪声$\xi(t)$和信号谱$P(\lambda)$满足什么条件能存在$x_0\in H_\xi(\tau)$使得

$$\frac{|s_{x_0}|}{\sqrt{E|\xi_{x_0}|^2}}=(S/N)_\tau, \tag{3.221}$$

并要求出x_0对$\{\xi(t)|t\leqslant\tau\}$的显明表达式(或用$P(\lambda)$及$\xi(t)$的谱$F(\lambda)$表示)。

引理3.2 设$\mathrm{d}P\ll\mathrm{d}F$，$\frac{\mathrm{d}P}{\mathrm{d}F}\in L^2(\mathrm{d}F)$，则

$$x\in H_\xi\Longleftrightarrow x=\int_{-\pi}^{\pi}\Phi(\lambda)\mathrm{d}Z_\xi(\lambda)+\int_{-\pi}^{\pi}\Phi(\lambda)\mathrm{d}P(\lambda), \tag{3.222}$$

其中$\Phi(\lambda)\in L^2(\mathrm{d}F)$唯一由$x$所决定。$F(\lambda)$是$\xi(t)$的谱函数，$L^2(\mathrm{d}F)$为$\mathrm{d}F$测度所张成的$L^2$空间。

证明 先证必需性。设$x\in H_\xi$，由引理3.1知$x=\xi_x+s_x$，并且有$\|\xi_x-\xi_x^{(n)}\|^2\to0$，其中

$$\xi_x^{(n)}=\int_{-\pi}^{\pi}\left(\sum_{j=1}^{k_n}a_j^{(n)}\mathrm{e}^{\mathrm{i}\lambda t_j^{(n)}}\right)\mathrm{d}Z_\xi(\lambda). \tag{3.223}$$

由$L^2(\mathrm{d}F)$性质，令

$$\Phi(\lambda)=\lim_{n\to\infty}\sum_{j=1}^{k_n}\alpha_j^{(n)}e^{i\lambda t_j^{(n)}}\ (L^2(dF)),\tag{3.224}$$

则

$$\xi_x=\int_{-\pi}^{\pi}\Phi(\lambda)dZ_\xi(\lambda).\tag{3.225}$$

又由 $dP\ll dF$，则

$$s_x=\lim_{n\to\infty}\sum_{j=1}^{k_n}\alpha_j^{(n)}s(t_j^{(n)})=\lim_{n\to\infty}\sum_{j=1}^{k_n}\alpha_j^{(n)}\int_{-\pi}^{\pi}e^{i\lambda t_j^{(n)}}dP(\lambda)$$

$$=\lim_{n\to\infty}\int_{-\pi}^{\pi}\left(\sum_{j=1}^{k_n}\alpha_j^{(n)}e^{i\lambda t_j^{(n)}}\right)\frac{dP}{dF}dF.$$

由于$\frac{dP}{dF}\in L^2(dF)$，于是由 L^2 空间的弱收敛性(见[6])知

$$s_x=\int_{-\pi}^{\pi}\Phi(\lambda)\left(\frac{dP}{dF}\right)dF=\int_{-\pi}^{\pi}\Phi(\lambda)dP,\tag{3.226}$$

故(3.222)成立。

反之，若 $\Phi(\lambda)\in L^2(dF)$而(3.224)成立，由弱收敛性知(3.226)成立。

令

$$x_n=\sum_{j=1}^{k_n}\alpha_j^{(n)}(\xi(t_j^{(n)})+s(t_j^{(n)}))$$

则当 $n,m\to\infty$时，

$$\|x_n-x_m\|^2=\|\xi_x^{(n)}-\xi_x^{(m)}\|^2+|s_x^{(n)}-s_x^{(m)}|^2\to0.$$

可见$\{x_n\}$在 H_ξ 中是 Cauchy 列，即

$$x=\lim_n x_n\in H_\xi.\quad \blacksquare$$

以后把(3.222)写成

$$x=\int_{-\pi}^{\pi}\Phi(\lambda)\zeta(d\lambda).\tag{3.227}$$

引理3.3 设 $\mathrm{d}P \ll \mathrm{d}F$，$\dfrac{\mathrm{d}P}{\mathrm{d}F} \in L^2(\mathrm{d}F)$，则

$$x \in H_\xi(\tau) \Longleftrightarrow x = \int_{-\pi}^{\pi} \Phi(\lambda)\zeta(\mathrm{d}\lambda),\quad \Phi(\lambda) \in L_\tau^2(\mathrm{d}F).$$

这样的 Φ 由 x 唯一确定，其中

$$L_\tau^2(\mathrm{d}F) = \mathscr{L}\{\mathrm{e}^{\mathrm{i}\lambda t} | t \leqslant \tau\}.$$

证明 与引理3.2相类似，只需注意 t 限制在 $\leqslant \tau$ 之内。∎

定理3.12 设 $\mathrm{d}P \ll \mathrm{d}F$，$\dfrac{\mathrm{d}P}{\mathrm{d}F} \in L^2(\mathrm{d}F)$，则(3.221)的最优解必存在唯一，并且若令 $\Phi_0(\lambda)$ 为 $\overline{\left(\dfrac{\mathrm{d}P}{\mathrm{d}F}\right)}$ 在 $L_\tau^2(\mathrm{d}F)$ 中的投影，则

$$x_0 = \int_{-\pi}^{\pi} \Phi_0(\lambda)\zeta(\mathrm{d}\lambda). \tag{3.228}$$

此时

$$(S/N)_\tau = \sqrt{(\Phi_0, \Phi_0)_{L^2(\mathrm{d}F)}}\ .$$

证明 由引理3.3知(3.220)可改写为

$$\begin{aligned}
(S/N)_\tau &= \sup_{x \in H_\xi(\tau)} \frac{|s_x|}{\sqrt{E|\zeta_x|^2}} \\
&= \sup_{\Phi \in L_\tau^2(\mathrm{d}F)} \frac{\left|\int_{-\pi}^{\pi} \Phi(\lambda)\mathrm{d}P\right|}{\sqrt{E\left|\int_{-\pi}^{\pi} \Phi(\lambda)\mathrm{d}Z_\xi(\lambda)\right|^2}} \\
&= \sup_{\Phi \in L_\tau^2(\mathrm{d}F)} \frac{\left|\int_{-\pi}^{\pi} \Phi(\lambda)\left(\dfrac{\mathrm{d}P}{\mathrm{d}F}\right)\mathrm{d}F\right|}{\sqrt{\int_{-\pi}^{\pi} |\Phi(\lambda)|^2 \mathrm{d}F}} \\
&= \sup_{\Phi \in L_\tau^2(\mathrm{d}F)} \frac{\left|\left(\Phi, \overline{\dfrac{\mathrm{d}P}{\mathrm{d}F}}\right)\right|}{\sqrt{(\Phi, \Phi)}}.
\end{aligned} \tag{3.229}$$

由 Hilbert 空间的 Riesz 分解定理，对子空间 $L_\tau^2(\mathrm{d}F)$ 可把 $\overline{\left(\dfrac{\mathrm{d}P}{\mathrm{d}F}\right)}$

分解为

$$\left(\overline{\frac{\mathrm{d}P}{\mathrm{d}F}}\right)=\Phi_0\oplus\Phi_\perp, \tag{3.230}$$

其中 Φ_0 为投影于 $L^2_\tau(\mathrm{d}F)$的分量。从而

$$(S/N)_\tau=\sup_{\Phi\in L^2_\tau(\mathrm{d}F)}\frac{|(\Phi,\Phi_0)|}{\|\Phi\|}. \tag{3.231}$$

利用 Schwartz 不等式得

$$|(\Phi,\Phi_0)|\leqslant\|\Phi\|\|\Phi_0\|,$$

等号成立的充要条件为 $\Phi=a\Phi_0$，因此由(3.231)可得

$$\begin{aligned}(S/N)_\tau&=\sup_{\Phi\in L^2_\tau(\mathrm{d}F)}\frac{|(\Phi,\Phi_0)|}{\|\Phi\|}\\&\leqslant\sup_{\Phi\in L^2_\tau(\mathrm{d}F)}\frac{\|\Phi\|\|\Phi_0\|}{\|\Phi\|}\\&=\|\Phi_0\|,\end{aligned} \tag{3.232}$$

等号成立的充要条件为

$$x_0=a\left(\int_{-\pi}^{\pi}\Phi_0(\lambda)\zeta(\mathrm{d}\lambda)\right). \tag{3.233}$$

因唯一性可差一常数倍，可取 $a=1$。 ▎

推论1 在本定理条件下，

$$(S/N)_\tau\leqslant\sqrt{\int_{-\pi}^{\pi}\left|\frac{\mathrm{d}P}{\mathrm{d}F}\right|^2\mathrm{d}F}, \tag{3.234}$$

等号成立$\Longleftrightarrow\left(\overline{\frac{\mathrm{d}P}{\mathrm{d}F}}\right)\in L^2_\tau(\mathrm{d}F)$,

此时最优解的谱特征为

$$\Phi_0(\lambda)=\left(\overline{\frac{\mathrm{d}P}{\mathrm{d}F}}\right).$$

推论2 (Dwork[31]) 设 $\xi(t)$的谱函数绝对连续，其谱密度满足 $f^{-1}(\lambda)\in L_1(\mathrm{d}\lambda)$，信号 $s(t)$可表为

$$s(t)=\int_{-\pi}^{\pi}e^{i\lambda t}\Omega_s(\lambda)\mathrm{d}\lambda \tag{3.235}$$

且 $\Omega_s(\lambda)$为连续函数，则

$$(S/N)_\tau \leqslant \sqrt{\int_{-\pi}^{\pi} \frac{|\Omega_s(\lambda)|^2}{f(\lambda)} d\lambda}, \tag{3.236}$$

等号成立$\Longleftrightarrow$

$$\overline{\left(\frac{\Omega_s(\lambda)}{f(\lambda)}\right)} \in L_\tau^2(dF), \tag{3.237}$$

此时最优解为

$$\Phi_0(\lambda) = \overline{\left(\frac{\Omega_s(\lambda)}{f(\lambda)}\right)}. \tag{3.238}$$

证明 在所设的条件下有

$$F(\lambda) = \int_{-\pi}^{\lambda} f(\mu) d\mu, \quad P(\lambda) = \int_{-\pi}^{\lambda} \Omega_s(\mu) d\mu,$$

并有 $dP \ll d\lambda \ll dF$，从而

$$\overline{\left(\frac{dP}{dF}\right)} = \overline{\left(\frac{dP}{d\lambda}\frac{d\lambda}{dF}\right)} = \overline{\left(\frac{\Omega_s(\lambda)}{f(\lambda)}\right)}.$$

其余由推论 1 自明。 ▎

以下定理给出当噪声$\xi(t)$是正则序列条件下最优滤波的谱特征的表达式。

定理3.13 设 $dP \ll dF$，$\frac{dP}{dF} \in L^2(dF)$，$\xi(t)$是正则(见[13])平稳列，则最优解为

$$x_0 = \int_{-\pi}^{\pi} \Phi_0(\lambda) \zeta(d\lambda),$$

其中

$$\Phi_0(\lambda) = \sum_{k=-\tau}^{\infty} a_k \frac{e^{-ik\lambda}}{\Gamma_{\xi\xi}(e^{-i\lambda})}, \tag{3.239}$$

$$a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\left(\frac{dP}{dF}\right)} \Gamma_{\xi\xi}(e^{-i\lambda}) e^{ik\lambda} d\lambda \quad (k \geqslant -\tau), \tag{3.240}$$

$\Gamma_{\xi\xi}(z)$由(3.170)定义，并且

$$(\cup, N)_{\tau} = \left(\sum_{k=-\tau}^{\infty} |a_k|^2 \right)^{\frac{1}{2}}. \tag{3.241}$$

证明 因为$\overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)} \in L^2(\mathrm{d}F)$，故利用本节第一段的方法，令

$$y \triangleq \int_{-\pi}^{\pi} \overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)} \mathrm{d}Z_{\xi}(\lambda) = \sum_{k=-\infty}^{\infty} a_k \varepsilon(-k). \tag{3.242}$$

$$\begin{aligned} a_k &= (y, \varepsilon(-k)) \\ &= \left(\int_{-\pi}^{\pi} \overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)} \Gamma_{\xi\xi}(\mathrm{e}^{-\mathrm{i}\lambda}) \mathrm{d}Z_{\varepsilon}(\lambda), \quad \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}k\lambda} \mathrm{d}Z_{\varepsilon}(\lambda) \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)} \Gamma_{\xi\xi}(\mathrm{e}^{-\mathrm{i}\lambda}) \mathrm{e}^{\mathrm{i}k\lambda} \mathrm{d}\lambda, \end{aligned} \tag{3.243}$$

$$k = 0, \pm 1, \pm 2, \cdots,$$

其中$\{\varepsilon(k)\}$为$\xi(t)$的 Wold 序列(见(3.163))．由$\{\varepsilon(k)\}$的等时从属性知

$$\begin{aligned} y &= \sum_{k<-\tau} a_k \varepsilon(-k) \oplus \sum_{k \geq -\tau} a_k \varepsilon(-k) \\ &= y_{\perp} \oplus y_0, \end{aligned} \tag{3.244}$$

其中 $y_0 \in H_{\xi}(\tau)$，$y_{\perp} \perp H_{\xi}(\tau)$。由 H_{ξ} 与 $L^2(\mathrm{d}F)$的同构对应关系，(3.244)相当于将$\overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)}$对 $L^2_{\tau}(\mathrm{d}F)$作正交分解

$$\overline{\left(\frac{\mathrm{d}P}{\mathrm{d}F}\right)} = \Phi_0 \oplus \Phi_{\perp}.$$

显见

$$y_0 = \sum_{k \geq -\tau}^{\infty} a_k \varepsilon(-k) = \int_{-\pi}^{\pi} \Phi_0(\lambda) \mathrm{d}Z_{\xi}(\lambda),$$

$$\Phi_0(\lambda) = \sum_{k \geq -\tau}^{\infty} a_k \frac{\mathrm{e}^{-\mathrm{i}k\lambda}}{\Gamma_{\xi\xi}(\mathrm{e}^{-\mathrm{i}\lambda})} \quad (\text{见}(3.166)).$$

再由定理3.12知

$$(S/N)_\tau = \|\Phi_0\| = \|y_0\| = \left(\sum_{k\geq -\tau}^{\infty} |a_k|^2\right)^{\frac{1}{2}}. \quad \blacksquare$$

在[24]中还给出当$\xi(t)$是奇异的或一般非奇异(未必正则)场合下，极大信噪比滤波的结果。

以下定理概括了在工程技术中广为流传的North(见[32])的结果，它是在干扰为白噪声下的特殊结果。

定理3.14(North) 设 $\xi(t)$是具有绝对连续谱且其谱密度为常值的平稳列$f(\lambda)=N_0>0$，

$$s(t)=\int_{-\pi}^{\pi} e^{i\lambda t}\Omega_s(\lambda)d\lambda$$

其中$\Omega_s(\lambda)$为连续函数，则以下三个结论等价：

(1) $$(S/N)_\tau = \sqrt{\int_{-\pi}^{\pi}\frac{|\Omega_s(\lambda)|^2}{N_0}d\lambda};\tag{3.245}$$

(2) $\bar{\Omega}_s(\lambda)\in L_\tau^2(dF)$;

(3) $s(t)=0$，当$t>\tau$最优解

$$x_0 = c\sum_{k=-\infty}^{\tau}\bar{s}(k)\xi(k),\tag{3.246}$$

其中c为常数。

证明 不难看出满足本定理的白噪声$\xi(t)$是正则的正交列，而且Wold序列为

$$\varepsilon(t)=\frac{1}{\sqrt{2\pi N_0}}\xi(t).\tag{3.247}$$

令

$$y=\int_{-\pi}^{\pi}\left(\frac{dP}{dF}\right)dZ_\xi(\lambda) = \int_{-\pi}^{\pi}\left(\frac{\overline{\Omega_s(\lambda)}}{N_0}\right)dZ_\xi(\lambda)$$

$$= \sum_{k=-\infty}^{\infty} a_k \varepsilon(k), \tag{3.248}$$

其中

$$\begin{aligned} a_k &= (y,\ \varepsilon(k)) \\ &= \left(\frac{\bar{\Omega}_s(\lambda)}{N_0},\ \frac{e^{ik\lambda}}{\sqrt{2\pi N_0}}\right) \\ &= \int_{-\pi}^{\pi} \left(\frac{\bar{\Omega}_s(\lambda)}{N_0}\right) \frac{e^{-ik\lambda}}{\sqrt{2\pi N_0}} N_0 d\lambda \\ &= \frac{1}{\sqrt{2\pi N_0}} \overline{\left\{\int_{-\pi}^{\pi} \Omega_S(\lambda) e^{ik\lambda} d\lambda\right\}} \\ &= \frac{\overline{s(k)}}{\sqrt{2\pi N_0}}, \quad k = 0, \pm 1, \cdots. \end{aligned} \tag{3.249}$$

由(3.234)的推论1知：

$$(S/N)_\tau = \sqrt{\int_{-\pi}^{\pi} \left|\frac{dP}{dF}\right|^2 dF} \Longleftrightarrow \left(\frac{dP}{dF}\right) = \frac{\Omega_s(\lambda)}{N_0} \in L_\tau^2(dF).$$

又由于

$$\frac{\bar{\Omega}_s(\lambda)}{N_0} = \sum_{k=-\infty}^{\infty} \bar{s}(k) \frac{e^{ik\lambda}}{2\pi N_0},$$

故

$$\begin{aligned} &\frac{\bar{\Omega}_s(\lambda)}{N_0} \in L_\tau^2(dF) \equiv L_\tau^2(d\lambda) \\ &\Longleftrightarrow s(k) = 0, \text{ 当} k > \tau. \end{aligned} \tag{3.250}$$

可见定理的三条件等价。代入定理3.12的(3.228)得

$$\begin{aligned} x_0 &= \sum_{k=-\infty}^{\tau} \frac{\bar{s}(k)}{\sqrt{2\pi N_0}} \left(\int_{-\pi}^{\pi} e^{ik\lambda} dZ_\xi(\lambda) + \int_{-\pi}^{\pi} e^{ik\lambda} \Omega_s(\lambda) d\lambda\right) \\ &= C \sum_{k=-\infty}^{\tau} \bar{s}(k) \zeta(k). \quad \blacksquare \end{aligned}$$

此定理即为North匹配滤波，由定理可看出：

(1) $(S/N)\tau=\left(\int_{-\pi}^{\pi}\frac{|\Omega_s(\lambda)|^2}{N_0}d\lambda\right)^{\frac{1}{2}}\Longleftrightarrow s(t)=0,\qquad t>\tau.$

$$(3.251)$$

表明使 τ 时刻的极大信噪比达到极大并等于(3.251)必须在 τ 时刻以前将全部信号输入于滤波器。这一要求在雷达中往往可以做到，因为雷达发出的信号往往是有间歇的一串串脉冲信号。但在许多工程技术问题中，这一要求虽然做不到，但并不意味着不能进行极大信噪比滤波，只是方法比较复杂一点，需用定理3.13的结果，并且 $(S/N)_\tau$ 达不到(3.234)的等式成立。

(2) North滤波即对信号波形进行匹配滤波，本质上，由(3.246)看就是一个"相关器"，也就是对收到的 $\xi(t)$ 序列用已知信号形式 $S(t)$ 进行"相关"滤波(运算)，因为噪声是白噪声。

更一般的极大信噪比滤波可参看[33]，[34]。

习　题

1. 设 x_t 为MA(2)模型

$$x_t=12\varepsilon_t-10\varepsilon_{t-1}+2\varepsilon_{t-2},$$

求：用Яглом方法的预报谱特征(一步预报与二步预报)。

2. 用时域方法写出上述MA(2)的向量预报公式。

3. 已知 x_t 为AR(3)模型

$$x_t-\frac{26}{24}x_{t-1}+\frac{9}{24}x_{t-2}-\frac{1}{24}x_{t-3}=2\varepsilon_t,$$

求对 $\tau=3,4$ 的预报公式

4. 已知 x_t 为ARMA(1,1)模型

$$x_t-\frac{1}{2}x_{t-1}=2\varepsilon_t-\varepsilon_{t-1},$$

用谱域和时域方法给出 $\tau=1,2$ 的预报公式。

5. 设 x_t 是正则的平稳列，则 x_t 是AR(p)的充分必要条件

是

$$E[x_{t+\tau}|H_x(t)]=a_1^{(\tau)}x_t+\cdots+a_p^{(\tau)}x_{t-p+1}.$$

6. 设 x_t 是平稳列，则下列三个命题等价

(a) $E[x_{t+\tau}|H_x(t)]=a_\tau x_t,\ \forall\tau>0,\ 0<|a_\tau|<1$;

(b) x_t 是 AR(1);

(c) 相关函数 $R_x(\tau)=a\rho^\tau,\ \tau\geqslant 0,\ a>0,\ 0<|\rho|\leqslant 1$.

7. 设 x_t 为 AR(2) 模型 $\Phi(U)x_t=\varepsilon_t$，而 $\Phi(z)=1-\varphi_1 z-\varphi_2 z^2$ 有二个复根 z_1^{-1},z_2^{-1}，其中 $z_1=\rho(\cos\gamma+i\sin\gamma)$，$|z_1|<1$。试证已知$\{x_s|s\leqslant t\}$之下对 $t+\tau$ 的预报公式可写为

$$E[x_{t+\tau}|H_x(t)]=x_t\rho^\tau\frac{\sin(\tau+1)\gamma}{\sin\gamma}-x_{t-1}\rho^{\tau+1}\frac{\sin\tau\gamma}{\sin\gamma}.$$

8. 在定理3.3中，当$p\geqslant q$的预报问题需确定一个 $p-1$ 次的多项式以 $a_1,a_2,\cdots,a_p$ 为根，实质上是用待定系数法，求解一个 p 个未知数的线性方程组，试证明其系数行列式是一个 Vandermonde 行列式，其值不等于零；而当$p<q$，则是一个广义的Vandermonde 行列式，其值也不为零。

9. 在极大信噪比滤波问题中，设 $dP\ll dF$，$\frac{dP}{dF}\in L^2(dF)$，而 $\xi(t)$ 为奇异的平稳列，则

$$(S/N)\tau=\left(\int_{-\pi}^{\pi}\left|\frac{dP}{dF}\right|^2 dF\right)^{\frac{1}{2}}$$

且最优滤波为

$$x_0=\int_{-\pi}^{\pi}\left(\overline{\frac{dP}{dF}}\right)\zeta(d\lambda).$$

10. 在极大信噪比滤波问题中，设 $dP\ll dF$，$\frac{dP}{dF}\in L^2(dF)$，而 $\xi(t)$为非奇异的平稳列，则最优滤波谱特征为

$$\Phi_0=\hat{\Phi}_0+\left(\overline{\frac{dP}{dF}}\right)\chi_A,$$

其中集A使

$$\int_{-\pi}^{\pi} \chi_A \mathrm{d}\lambda = \int_{-\pi}^{\pi} \chi_{\overline{A}} \mathrm{d}\, F_{ss} = 0,$$

而 $\mathrm{d}F_{ss}$ 是对应于 $\mathrm{d}F(\lambda)$的奇异测度部分，Φ_0是$\left(\overline{\frac{\mathrm{d}P}{\mathrm{d}F}}\right)\chi_A$对正则谱 $F_{rr}(\lambda)$按定理3.13得到的最优解的谱特征，此时

$$(S/N)_r = \left(\int_{-\pi}^{\pi}\left(|\Phi_0|^2 + \left|\frac{\mathrm{d}P}{\mathrm{d}F}\right|^2 \chi_A\right)\mathrm{d}F\right)^{\frac{1}{2}}.$$

第四章 谱估计的参数方法

§1 引 言

1° 谱估计问题的提法和意义

在许多工程技术、自然科学、社会经济等领域中，根据观测数据来分析研究对象的周期性或能量在频域上的分布具有十分重要的意义。数学上通常的提法是：设研究的对象是平稳过程$x(t)$，通过实验观测，得到它的记录$\{x(s),\ s\in T\}$，希望给出一个统计量$\theta_\lambda\{x(s),\ s\in T\}$，$\lambda\in\Pi$，用来估计过程$x(t)$的谱，并且该估计量$\theta_\lambda$在一些统计或概率的性质上具有某些优良性。具体一些说，比如$x(t)$是平稳序列，观测样本为$x(s)$，$s=1,2,\cdots,N$，希望构造一个$\theta_\lambda\{x_1,\cdots,x_N\}$作为$x(t)$的谱估计。由于平稳列的谱可分为以下三种：

A．$dF_x\ll d\lambda$，则$\theta_\lambda\{x_1,\cdots,x_N\}$是用来估计谱密度$f_x(\lambda)$；

B．$F_x(\lambda)$是有限个跳跃的离散谱，这时的$\theta_\lambda\{x_1,\cdots,x_N\}$是用于估计诸跳跃点$\{\lambda_k\}$和阶跃高度$\{\Delta F(\lambda_k)\}$；

C．$F_x(\lambda)$是具有绝对连续部分和跳跃部分的混合谱。因此$\theta_\lambda\{x_1,\cdots,x_N\}$应该兼顾二部分的估计。

2° 谱估计方法中的两类重要途径

和经典统计学中的估值理论相类似，谱估计方法中也可类比为参数方法与非参数方法。所谓参数方法（见[37]）是指：假定$x(t)$（研究对象）具有有限个模型参数，如ARMA(p,q)具有$p+q+1$个参数，因而$x(t)$的谱密度$f_x(\lambda)$也依赖于有限个参数，表为$f_x(\lambda;\ \theta)$，θ可能限定在某个合理的范围$\theta\in\Theta$。假若测得$x(t)$的

观测值 $x_1, x_2, \cdots, x_N$，希望由它们给出 θ 的合理估计 $\hat{\theta}$，并用 $f_N(\lambda, \hat{\theta})$ 来估计 $f_x(\lambda)$。所谓非参数方法是不假定 $x(t)$ 是具有有限参数的模型，因而希望直接由样本 $x_1, x_2, \cdots, x_N$ 给出一个 $f_x(\lambda)$ 的估计量 $\hat{f}_N(\lambda)$。

值得注意的是近二十年间发展起来的方法是：不管 $x(t)$ 原本是否是具有有限参数的模型，由 $x(t)$ 的观测 $x_1, x_2, \cdots, x_N$ 出发，给出对 $x(t)$ 的一种具有有限参数模型 $\tilde{x}(t, \theta)$ 的拟合，而后者的谱密度 $f_N(\lambda; \hat{\theta})$ 是 $x(t)$ 谱密度 $f_x(\lambda)$ 的一种好的估计量。我们将把这类方法归为"参数方法"之中。

3° 参数谱估计方法与模型拟合的同一性

如果对 $x(t)$ 的实际认识只有 N 个样本 $x_1, x_2, \cdots, x_N$，如果我们能够对 $x(t)$ 给它一种合理的参数模型的拟合 $\tilde{x}(t, \hat{\theta})$，则对 $f_x(\lambda)$ 的谱估计就往往可以利用模型的知识给它一种参数谱估计 $f_N(\lambda; \hat{\theta})$。比如对 $x(t)$ 给它 ARMA(p, q) 的模型拟合是 $\tilde{x}(t, \varphi, \theta)$，则由第二章的知识我们知道这时

$$f_N(\lambda; \varphi, \theta) = \frac{1}{2\pi} |\Theta(e^{-i\lambda}) / \Phi(e^{-i\lambda})|^2$$

就是一种估计。

本章将介绍 AR, MA, 长阶 AR 和 Bloomfield 参数模型的谱估计方法，并介绍一些实用的判阶方法。

§2 信息准则下对实测数据的模型拟合和谱估计

1° 信息论中的某些基本知识

为了让读者了解近代时间序列分析的某些特点并扩大知识面，以下介绍一些信息理论的基本知识。

首先要指出：从统计观点看，一个事件 A 的发生如果能给人们带来了信息，则它应该是随机事件——即人们无法完全预料的

事件．显见，一件为人们所完全预料(已成为常识)的事件就不再会给人们带来信息．假定 A,B 是随机事件，$P(A)>P(B)$，人们从常识中易知：小概率事件的发生给人们带来的信息显然比大概率要多，即 B 事件的信息比 A 要多．据此，Shannon（见[38]）等人建议用数量

$$-\log P(A) \tag{4.1}$$

来刻画 A 给人们带来的信息量，并建议 log 以 2 为底，这样相当于掷一次硬币的结果给人们带来的信息量为一个信息单位，称为 bit．

如果考查一个随机变量 ξ(离散型)，它的分布为

$$\xi \sim \begin{pmatrix} A_1 & A_2 & \cdots & A_n \\ P_1 & P_2 & \cdots & P_n \end{pmatrix}. \tag{4.2}$$

显见，ξ 发生的随机事件可为 $A_1,\cdots,A_n$，平均说来（依概率平均），每发生一事件给人们带来的信息量为

$$H(\xi) = -\sum_{i=1}^{n} p_i \log p_i, \tag{4.3}$$

称为“熵”(ENTROPY)．

如果 ξ 是连续型的，Shannon 建议用类似的方法定义 ξ 的熵：

$$H(\xi) = -\int_{R_1} p_\xi(x)\log p_\xi(x)\mathrm{d}x. \tag{4.4}$$

从微积分常识知：当然不是所有分布密度都可使得(4.4)存在(参看本章习题)．

如果 $E\xi, D(\xi)$ 皆存在，并且 $D(\xi)=\sigma^2$ 给定，则以下可看到：使(4.4)达极大的分布正好是正态分布(方差为 σ^2)．

如果 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)$ 有联合分布，以密度为例，是 $p_\xi(x)$，$x\in \boldsymbol{R}_n$，则称

$$H(\xi) = H(\xi_1,\xi_2,\cdots,\xi_n) = -\int_{R_n}\cdots\int p_\xi(x)\log p_\xi(x)\mathrm{d}x \tag{4.5}$$

为ξ的熵。

如果ξ_t是一个随机过程，以序列为例，先对任意的正整数n，讨论$H(\xi_1,\xi_2,\cdots,\xi_n)$的熵。如果

$$H(\xi_t)=\lim_{n\to\infty}\frac{1}{n}H(\xi_1,\xi_2,\cdots,\xi_n) \tag{4.6}$$

存在，就称$H(\xi_t)$为ξ_t的"熵率"。

为了得到一维、多维乃至过程的熵的某些具体表达式。我们引入几个定理(可参看[39])。

定义4.1(Kullback-Leibler 信息量)　设$p_1(x),p_2(x)$是共同定义于$E,E\subset \boldsymbol{R}_1$上的概率密度函数，并且$p_1(x)p_2(x)>0$，$x\in E$，则称

$$I(p_1,p_2)=\int_E p_1(x)\log\frac{p_1(x)}{p_2(x)}\,\mathrm{d}x \tag{4.7}$$

为(p_1,p_2)的K-L信息量。

定理4.1

$$I(p_1,p_2)\geqslant 0, \tag{4.8}$$

等号成立的充分必要条件是

$$p_1(x)=p_2(x)\quad \text{a.e.}(\mathrm{d}x)。 \tag{4.9}$$

证明　首先指出，对$y>0$，恒有

$$\log y\leqslant y-1 \tag{4.10}$$

成立，等号成立的充要条件是$y=1$。事实上，令

$$f(y)=y-1-\log y,$$

则

$$f'(y)=1-\frac{1}{y},\qquad f'(1)=0。$$

$$f''(y)=\frac{1}{y^2}>0\ (对y>0)。$$

表明$f(y)$只有一个极小值点，即$y=1$。而$f(1)=0$，故$f(y)\geqslant 0$，

等号成立的充分必要条件是$y=1$。

将上述结果用于 K-L 信息量的表达式(4.7)可得

$$\int_E p_1(x)\log\frac{p_2(x)}{p_1(x)}\mathrm{d}x\leqslant\int_E p_1(x)\left(\frac{p_2(x)}{p_1(x)}-1\right)\mathrm{d}x$$

$$=\int_E(p_2(x)-p_1(x))\mathrm{d}x=0。\tag{4.11}$$

可见

$$I(p_1,p_2)=-\int_E p_1(x)\log\frac{p_2(x)}{p_1(x)}\mathrm{d}x\geqslant0,$$

等号成立的充分必要条件是

$$\frac{p_2(x)}{p_1(x)}=1\quad \text{a.e.}(\mathrm{d}x). \ \|$$

定理4.2 设$\xi=(\xi_1,\xi_2,\cdots,\xi_n)$是$n$维随机变量，有概率密度$p_\xi(x)>0$，$x\in\boldsymbol{R}_n$，并且其协方差阵

$$[\mathrm{Cov}(\xi_i,\xi_j)]=\begin{bmatrix}\sigma_{11}&\sigma_{12}&\cdots&\sigma_{1n}\\ \vdots&\ddots&&\vdots\\ \sigma_{n1}&\sigma_{n2}&\cdots&\sigma_{nn}\end{bmatrix}\triangleq\boldsymbol{\Gamma}_n>0\tag{4.12}$$

给定，则使熵$H(\xi)$(见(4.5))达极大的充分必要条件是ξ为正态分布，此时

$$H_{\max}(\xi)=\log((2\pi e)^{\frac{n}{2}}(\det\boldsymbol{\Gamma}_n)^{\frac{1}{2}}).\tag{4.13}$$

证明 首先指出有关 K-L 信息量的定理4.1，在多维场合下也得到完全类似的结果。取$p_2(\boldsymbol{x})$为$N(O,\boldsymbol{\Gamma}_n)$的正态分布密度，于是，由(4.11)知

$$\int_{R_n}p_1(\boldsymbol{x})\log(p_2(\boldsymbol{x})/p_1(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}$$

$$=\int_{R_n}p_1(\boldsymbol{x})\log p_2(\boldsymbol{x})\mathrm{d}\boldsymbol{x}-\int_{R_n}p_1(\boldsymbol{x})\log p_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\leqslant0$$

或

$$-\int_{R_n} p_1(x)\log p_1(x)\mathrm{d}x \leqslant -\int_{R_n} p_1(x)\log p_2(x)\mathrm{d}x$$

$$= -\int_{R_n} p_1(x)\left[\log\left(\left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\frac{1}{(\det\Gamma_n)^{\frac{1}{2}}}e^{-\frac{1}{2}x\Gamma_n^{-1}x'}\right)\right]\mathrm{d}x$$

$$= \log\left((2\pi)^{\frac{n}{2}}(\det\Gamma_n)^{\frac{1}{2}}\right) + \frac{1}{2}\int_{R_n} p_1(x)x\Gamma_n^{-1}x'\mathrm{d}x. \quad (4.14)$$

利用代数知识：

$$x\Gamma_n^{-1}x' = \mathrm{tr}(\Gamma_n^{-1}x'x), \quad (4.15)$$

(4.14)可写为

$$-\int_{R_n} p_1(x)\log p_1(x)\mathrm{d}x$$

$$= \log\left((2\pi)^{\frac{n}{2}}(\det\Gamma_n)^{\frac{1}{2}}\right) + \frac{1}{2}\mathrm{tr}\left(\int_{R_n}(\Gamma_n^{-1}x'x)p_1(x)\mathrm{d}x\right)$$

$$= \log\left((2\pi)^{\frac{n}{2}}(\det\Gamma_n)^{\frac{1}{2}}\right) + \frac{1}{2}\mathrm{tr}\left(\Gamma_n^{-1}\int_{R_n}x'xp_1(x)\mathrm{d}x\right). \quad (4.16)$$

由于假定 $p_1(x)$ 的协方差阵为(4.12)，故

$$\Gamma_n^{-1}\int_{R_n}x'xp_1(x)\mathrm{d}x$$

$$= \Gamma_n^{-1}\left(\int_{R_n}x_ix_jp_1(x)\mathrm{d}x\right)_{1\leqslant i,j\leqslant n}$$

$$= I_n. \quad (4.17)$$

将(4.17)代入(4.16)即可得

$$H(\xi_1) \leqslant \log\left((2\pi)^{\frac{n}{2}}(\det\Gamma_n)^{\frac{1}{2}}\right) + \frac{1}{2}n$$

$$= \log\left((2\pi e)^{\frac{n}{2}}(\det\Gamma_n)^{\frac{1}{2}}\right). \quad (4.18)$$

等号成立的充分必要条件是

$$p_1(\boldsymbol{x})=p_2(\boldsymbol{x})\sim N(0,\ \boldsymbol{\Gamma}_n). \quad \blacksquare$$

设想 x_t 是一个正态的AR(p) 序列，以下来计算一下它的熵率 $H(x_t)$：

设$(\varphi_1,\varphi_2,\cdots,\varphi_p,\theta_0)$是 x_t 的 AR 模型参数，由第二章的知识知，它满足如下的 Yule-Walkor 方程

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_p \\ R_1 & R_0 & \cdots & R_{p-1} \\ \cdots & \cdots & \cdots & \cdots \\ R_p & R_{p-1} & \cdots & R_0 \end{bmatrix}\begin{bmatrix} 1 \\ \varphi_1 \\ \vdots \\ \varphi_p \end{bmatrix}=\begin{bmatrix} \theta_0^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

简记为

$$\boldsymbol{R}_{p+1}(1,\varphi_1,\cdots,\varphi_p)'=(\theta_0^2,0,\cdots,0)'. \tag{4.19}$$

由线性代数中的克拉姆法则知

$$1=\frac{\begin{vmatrix} \theta_0^2 & R_1 \quad \cdots \quad R_p \\ 0 & \\ \vdots & \boldsymbol{R}_p \\ 0 & \end{vmatrix}}{|\boldsymbol{R}_{p+1}|}=\frac{\theta_0^2\det\boldsymbol{R}_p}{\det\boldsymbol{R}_{p+1}},$$

即

$$\theta_0^2=c_0^2=\det\boldsymbol{R}_{p+1}/\det\boldsymbol{R}_p. \tag{4.20}$$

不难看出对一般的$N\geqslant p$皆有

$$c_0^2=\det\boldsymbol{R}_{N+1}/\det\boldsymbol{R}_N. \tag{4.21}$$

可见

$$\begin{aligned}\det\boldsymbol{R}_N&=c_0^2\det\boldsymbol{R}_{N-1}=\cdots\\&=(c_0^2)^{N-p}\det\boldsymbol{R}_p.\end{aligned}$$

于是由(4.18)得

$$\begin{aligned}H_N(x)&=\log((2\pi e)^{\frac{N}{2}}(\det\boldsymbol{R}_N)^{\frac{1}{2}})\\&=\frac{N}{2}\log(2\pi e)+\frac{1}{2}\log(\det\boldsymbol{R}_N)\end{aligned}$$

$$= \frac{N}{2}\log(2\pi e) + \frac{1}{2}\log((c_0^2)^{N-p}\det R_p)$$

$$= \frac{N}{2}\log(2\pi e) + \frac{N-p}{2}\log(c_0^2) + \frac{1}{2}\log(\det R_p).$$

可见

$$H(x_t) = \lim_{N\to\infty}\frac{1}{N}H_N(x)$$

$$= \frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(c_0^2). \qquad (4.22)$$

c_0^2 由第二章知是 x_t 的一步预测误差。

以上虽然是对 AR 型来证的，事实上一般的正态平稳序列皆有以上的类似结果(见[40])。

定理4.3 设 x_t 是 ARMA 模型，$f_x(\lambda)$是它的谱密度，则一步预测误差为

$$c_0 = \sqrt{2\pi}\,\exp\left\{\frac{1}{4\pi}\int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda\right\}. \qquad (4.23)$$

证明

$$f_x(\lambda) = \frac{1}{2\pi}|\Gamma_x(e^{-i\lambda})|^2 = \frac{1}{2\pi}\left|\frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})}\right|^2. \qquad (4.24)$$

$$\Gamma_x(z) = \frac{\Theta(z)}{\Phi(z)}$$

是$|z|\leqslant 1$ 内的解析函数，用幅角表示有

$$\log\Gamma_x(z) = \log|\Gamma_x(z)| + i\Psi(z),$$

因而$\log|\Gamma_x(z)|$是解析函数的实部，是调和函数。由平均值定理

$$\log|\Gamma_x(0)| = \frac{1}{2\pi}\int_{\Pi}\log|\Gamma_x(e^{-i\lambda})|\mathrm{d}\lambda$$

$$= \frac{1}{2\pi}\int_{\Pi}\log(\sqrt{2\pi f_x(\lambda)}\,)\mathrm{d}\lambda \text{（由(4.24)）}$$

$$= \log\sqrt{2\pi} + \frac{1}{4\pi}\int_{\Pi} \log f_x(\lambda)\mathrm{d}\lambda,$$

即

$$|\Gamma_x(0)| = e^{\log\sqrt{2\pi}} \exp\left\{\frac{1}{4\pi}\int_{\Pi} \log f_x(\lambda)\mathrm{d}\lambda\right\}$$

$$= \sqrt{2\pi}\exp\left\{\frac{1}{4\pi}\int_{\Pi} \log f_x(\lambda)\mathrm{d}\lambda\right\}. \qquad (4.25)$$

而 $c_0 = |\Gamma_x(0)|$，定理得证。 ▌

将(4.23)代入(4.22)知

$$H(x_t) = \frac{1}{2}\log(2\pi e) + \frac{1}{2}\log\left(\sqrt{2\pi}\exp\left(\frac{1}{4\pi}\int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda\right)\right)^2,$$

$$H(x_t) = \frac{1}{2} + \log(2\pi) + \frac{1}{4\pi}\int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda \qquad (4.26)$$

或

$$H(x_t) = C + \frac{1}{4\pi}\int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda. \qquad (4.27)$$

对正态平稳列，其谱密度 $f_x(\lambda)$满足 $\ln f_x(\lambda) \in L_1(\mathrm{d}\lambda)$，则(4.27)成立。

由(4.27)，(4.23)及(4.22)知：随机序列的熵率和一步预测误差及

$$I(f) = \int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda \qquad (4.28)$$

称为谱熵——三者之间是具有密切关系的。要熵率大$\Longleftrightarrow$谱熵 $I(f)$要大$\Longleftrightarrow$一步预测误差 c_0^2 要大。为什么说一步误差 c_0^2 大反倒信息量（熵率）大呢？熵率的直观意义是随机序列$\{x_t\}$每一个 x_t 的出现给人们带来的平均信息。显然，如果是能够准确预报的，则得到一段序列之后就可以知道"以后将发生什么"，必然熵率是很低的。

2° 最大一步预测误差准则下的模型拟合和谱估计

这一节里将介绍从平稳序列的一部分观测值如何对原过程进行模型拟合和谱估计的方法。先介绍在最大一步预测误差准则下如何由一部分相关函数值$R_0,R_1,\cdots,R_p$扩充成$\tilde{R}_0,\tilde{R}_1,\cdots,\tilde{R}_n,\cdots$而

$$\tilde{R}_k = R_k, \quad 0\leqslant k\leqslant p, \tag{4.29}$$

且$\{\tilde{R}_k\}$还是某个平稳列的相关函数。

定理4.4（存在性） 设$R_0,R_1,\cdots,R_p$是某个平稳列x_t的相关函数，

$$\boldsymbol{\Gamma}_{p+1}=\begin{bmatrix} R_0 & R_1 & \cdots & R_p \\ \vdots & \ddots & & \vdots \\ R_p & R_{p-1} & \cdots & R_0 \end{bmatrix}>0, \tag{4.30}$$

则存在AR(p)序列y_t，满足

$$R_y(k)=R_k, \quad 0\leqslant k\leqslant p, \tag{4.31}$$

y_t的模型参数$(\theta_0,\varphi_1,\cdots,\varphi_p)$由 Yule-Walker 方程

$$\boldsymbol{\Gamma}_{p+1}\begin{bmatrix} 1 \\ \varphi_1 \\ \vdots \\ \varphi_p \end{bmatrix}=\begin{bmatrix} \theta_0^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.32}$$

确定。

证明 由(4.20)知

$$\det\boldsymbol{\Gamma}_{p+1}=(\det\boldsymbol{\Gamma}_p)\theta_0^2>0.$$

由(4.30)知$\theta_0^2>0$。

令

$$R_k=\begin{cases} R_k, & 0\leqslant k\leqslant p; \\ -\sum_{n=1}^{p}\varphi_k R_{k-n}, & p<n. \end{cases} \tag{4.34}$$

用归纳法可以证明对$L\geqslant 1$有

$$\det\tilde{\Gamma}_{p+L+1}=(\det\tilde{\Gamma}_{p+L})\theta_0^2>0, \tag{4.35}$$

其中

$$\tilde{\Gamma}_{p+L+1}=\begin{bmatrix} \Gamma_p & \tilde{R}_p & \tilde{R}_{p+1} & \cdots & \tilde{R}_{p+L} \\ & & & & \vdots \\ & & & & \vdots \\ & & & & R_1 \\ \tilde{R}_{p+L} & \tilde{R}_{p+L-1} & \cdots & & R_0 \end{bmatrix} \tag{4.36}$$

$$=\begin{bmatrix} \tilde{\Gamma}_{p+L} & \tilde{R}_{p+L} \\ & \vdots \\ & \vdots \\ \tilde{R}_{p+L}\ \tilde{R}_{p+L-1}\ \cdots & R_0 \end{bmatrix}. \tag{4.37}$$

事实上，由分块行列式关系知

$$\det\tilde{\Gamma}_{p+L+1}=(\det\tilde{\Gamma}_{p+L})(R_0-\gamma'\tilde{\Gamma}_{p+L}^{-1}\gamma), \tag{4.38}$$

其中

$$\gamma'=(\tilde{R}_1,\tilde{R}_2,\cdots,\tilde{R}_{p+L}).$$

由(4.34)知

$$\tilde{R}_0-\gamma'\Gamma_{p+L}^{-1}\gamma=(\tilde{R}_0-(\tilde{R}_1,\cdots,\tilde{R}_{p+L})\begin{pmatrix} -\varphi_1 \\ \vdots \\ -\varphi_p \\ 0 \end{pmatrix}$$

$$=R_0+\sum_{l=1}^{p}\varphi_l R_l=\theta_0^2>0.$$

由(4.35)左边的等式知 $\det\Gamma_{p+L+1}>0$ $(\forall L\geqslant 1)$，可见

$$R_0,\ R_1,\ \cdots,\ R_p,\ R_{p+1},\ \cdots$$

是非负定数列。令

$$a_t\equiv 0,\quad t=0,\pm 1,\cdots;$$

$$B_n=(R_{(k-j)})_{1\leqslant k,\ j\leqslant n}>0,\quad \forall n>0,$$

则由第一章的 Gauss 过程存在定理(定理1.2)知必存在实Gauss 序列(平稳)y_t 以$\{\tilde{R}_k\}$为相关函数

$$R_y(k)=\begin{cases} R_k, & 0\leqslant k\leqslant p; \\ R_k, & p<k. \end{cases} \tag{4.39}$$

又由(4.34)知有

$$\Phi(U)R_y(n)=0, \quad n\geqslant 1, \tag{4.40}$$

其中

$$\Phi(z)=\sum_{k=0}^{p}\varphi_k z^k.$$

由(2.149)(定理 2.14)的性质知：y_t 必为AR(p)序列，并且

$$\Phi(z)\neq 0, \quad |z|\leqslant 1. \ \blacksquare \tag{4.40$'$}$$

定理4.5(极大性) 设 $R_0,R_1,\cdots,R_p$ 是给定的数列，它使得(4.30)成立，考虑一切平稳列$\{y_n\}$，它的相关函数 $R_y(k)$满足

$$R_y(k)=R_k, \quad 0\leqslant k\leqslant p,$$

则使$\{y_n\}$的一步预测误差 $c_0^{(y)}$ 达极大的充分必要条件是$\{y_n\}$为AR(p)模型。

证明 先证必需性.设$\{y_n\}$使 $c_0^{(y)}$ 达极大,即若有平稳列$\{z_n\}$其协方差也满足

$$R_z(k)=R_k, \quad 0\leqslant k\leqslant p, \tag{4.41}$$

则相应的预测误差应有

$$c_0^{(z)}\leqslant c_0^{(y)}. \tag{4.42}$$

由定理 4.4 的存在性，必可找到AR(p)序列$\{z_n\}$使

$$c_0^{(z)}=\theta_0^{(z)}=\left\|z_t-\sum_{k=1}^{p}\varphi_k^{(p)}z_{t-k}\right\|, \tag{4.43}$$

其中

$$\varphi_k^{(p)}=-\varphi_k, \quad k=1,2,\cdots,p.$$

取$\{-\varphi_k^{(p)}\}$为系数，应有

$$\left\|y_t-\sum_{k=1}^{p}\varphi_k^{(p)}y_{t-k}\right\|\geqslant\|y_t-\underset{H_y(t-1)}{\mathrm{Proj}}(y_t)\|=c_0^{(y)}. \tag{4.44}$$

然而(4.44)的左边

$$\left\| y_t - \sum_{k=1}^{p} \varphi_k^{(p)} y_{t-k} \right\|^2 = R_y(0) + \sum_{k=1}^{p} \varphi_k R_y(k)$$

$$= R_0 + \sum_{k=1}^{p} \varphi_k R_k = (\theta_0^{(x)})^2, \qquad (4.45)$$

可见

$$c_0^{(x)} = \theta_0^{(x)} \geqslant c_0^{(y)}.$$

又由(4.42)得

$$c_0^{(y)} = \left\| y_t - \sum_{k=1}^{p} \varphi_k^{(p)} y_{t-k} \right\|. \qquad (4.46)$$

表明

$$\underset{H_y(t-1)}{\text{Proj}}(y_t) = \underset{L[y_{t-1},\cdots,y_{t-p}]}{\text{Proj}}(y_t) = \sum_{k=1}^{p} \varphi_k^{(p)} y_{t-k}.$$

当$M > p$时，由$L[y_{t-1},\cdots,y_{t-M}] \subset H_y(t-1)$知

$$\underset{L[y_{t-1},\cdots,y_{t-M}]}{\text{Proj}}(y_t) = \underset{L[y_{t-1},\cdots,y_{t-p}]}{\text{Proj}}(y_t)$$

或

$$\sum_{k=1}^{M} \varphi_k^{(M)} y_{t-k} = \sum_{k=1}^{p} \varphi_k^{(p)} y_{t-k} \text{ (见定义2.4)}.$$

由投影的唯一性，$\varphi_M^{(M)} = 0$ $(M > p)$。可见$\{y_t\}$的偏相关系数$\{\varphi_k^{(k)}\}$是截尾的，由第二章定理2.14知y_t为AR(p)型。

再证充分性。假若$\{y_t\}$是AR(p)，则其系数$(\theta_0, \varphi_1, \cdots, \varphi_p)$必满足

$$\Gamma_{p+1}\begin{bmatrix} 1 \\ \varphi_1 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \theta_0^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

所以

$$c_0^{(p)}=\theta_0=\left\|y_t-\sum_{k=1}^{p}\varphi_k^{(p)}y_{t-k}\right\| \quad (\varphi_k=-\varphi_k^{(p)})$$

对满足 $R_z(k)=R_k, 0\leqslant k\leqslant p$ 的其他平稳列 $\{z_t\}$，必有

$$\left\|z_t-\sum_{k=1}^{p}\varphi_k^{(p)}z_{t-k}\right\|^2=R_z(0)+\sum_{k=1}^{p}\varphi_k R_z(k)$$

$$=R_0+\sum_{k=1}^{p}\varphi_k R_k=\theta_0^2\geqslant(c_0^{(p)})^2.$$

表明 $c_0^{(p)}$ 达极大。 ▌

定理 4.4 告诉我们存在 AR(p)序列$\{y_t\}$，它的模型参数$(\theta_0, \varphi_1,\cdots,\varphi_p)$由$\Gamma_{p+1}$确定的 Yule-Walker 方程解出，$R_y(k)=R_k$，$0\leqslant k\leqslant p$。显然 y_t 的模型必为

$$\sum_{k=0}^{p}\varphi_k y_{t-k}=\theta_0\varepsilon_t, \tag{4.47}$$

并且

$$y_t=\int_{\Pi}e^{i\lambda t}\frac{\theta_0}{\sum_{k=0}^{p}\varphi_k e^{-ik\lambda}}dZ_\varepsilon(\lambda), \tag{4.48}$$

因而$\{R_y(k)\}$完全由$(\theta_0,\varphi_1,\cdots,\varphi_p)$或$\{R_0,R_1,\cdots,R_p\}$决定。定理 4.5告诉我们，由(4.47)确定的$\{y_t\}$它的 $c_0^{(p)}$ 是最大的。

以上的讨论可以说是在一步预测误差达极大的准则下如何对给定的有限个相关函数值 $R_0,R_1,\cdots,R_p$ 给予扩充，$\{R_y(k)\}$就是这种扩充。而 $\{y_t\}$ 的谱密度 $f_y(\lambda)$ 由第二章的知识知完全由$\{R_y(k)\}$或$(\theta_0,\varphi_1,\cdots,\varphi_p)$确定：

$$f_y(\lambda)=\frac{\theta_0^2}{2\pi\left|1+\sum_{k=1}^{p}\varphi_k e^{-ik\lambda}\right|^2}, \tag{4.49}$$

从而 $f_y(\lambda)$就可以合理地充当 $f_x(\lambda)$的估计。

由上一节的讨论知道，最大熵率或最大谱熵或最大一步预测误差都是同一的，因而(4.49)在工程技术书或文献上(见[40])称为是最大熵谱估计，由(4.47)确定的$\{y_t\}$就称为最大熵准则下对原过程x_t的拟合模型。

3° 联合熵最大准则下的谱估计和模型拟合

上一节中我们介绍了在最大一步预报误差（或谱熵或熵率）准则下，对部分相关函数值$\{R_k; 0\leqslant k\leqslant p\}$的扩充并指出这种扩充的结果必和AR模型相联系，这一工作自Burg(见[41])提出以来有多种途径进行探讨。作为方法上的学习我们再介绍联合熵最大准则下的拟合问题。

设$R_0, R_1, \cdots, R_p$是某个平稳序列的$p+1$个相关函数值，并且$\boldsymbol{R}_{p+1}>0$，求一平稳列x_t使得

(1) $R_x(k)=R_k,\quad 0\leqslant k\leqslant p$;

(2) $\forall$正整数$L>1$，$(x_1, x_2, \cdots, x_{p+L})$的联合熵

$$H(x_1, x_2, \cdots, x_{p+L}) = -\int_{R_{p+L}} p(\boldsymbol{x})\log p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \quad (4.50)$$

在所有满足(1)的平稳列中是最大的。

由定理4.2知，这种平稳列x_t必定在正态分布类中达到。由(4.13)式知(4.50)为

$$H(x_1, x_2, \cdots, x_{p+L}) = \log((2\pi e)^{\frac{p+L}{2}}(\det\boldsymbol{\Gamma}_{p+L})^{\frac{1}{2}}). \quad (4.51)$$

由(4.51)知，问题化为由$R_0, R_1, \cdots, R_p$出发，如何逐步扩充$\tilde{R}_{p+1}, \tilde{R}_{p+2}, \cdots, \tilde{R}_{p+L}$，使每一步都使$\det\boldsymbol{\Gamma}_{p+L}$达极大。

为此我们先证明如下引理：

引理4.1 设$\alpha_0, \alpha_1, \cdots, \alpha_{m+1}$为$m+2$个$m$维数值向量，$\gamma_0, \gamma_1, \cdots, \gamma_m$为常数，令

$$
f(x) = \det \begin{bmatrix} \gamma_0 & \boldsymbol{\alpha}_0 & x \\ \gamma_1 & \boldsymbol{\alpha}_1 & \gamma_m \\ \vdots & \vdots & \vdots \\ \gamma_m & \boldsymbol{\alpha}_m & \gamma_1 \\ x & \boldsymbol{\alpha}_{m+1} & \gamma_0 \end{bmatrix}_{(m+2)\times(m+2)}, \tag{4.52}
$$

则

$$
\frac{\mathrm{d}f(x)}{\mathrm{d}x} = (-1)^{m-3}\left[\det\begin{bmatrix} \gamma_1 & \boldsymbol{\alpha}_1 \\ \vdots & \vdots \\ \gamma_m & \boldsymbol{\alpha}_m \\ x & \boldsymbol{\alpha}_{m+1} \end{bmatrix} + \det\begin{bmatrix} \boldsymbol{\alpha}_0 & x \\ \boldsymbol{\alpha}_1 & \gamma_m \\ \vdots & \vdots \\ \boldsymbol{\alpha}_m & \gamma_1 \end{bmatrix}\right]. \tag{4.53}
$$

证明 对(4.52)行列式可按第一列展开:

$$
f(x) = \gamma_0 \begin{vmatrix} \boldsymbol{\alpha}_1 & \gamma_m \\ \vdots & \vdots \\ \boldsymbol{\alpha}_m & \gamma_1 \\ \boldsymbol{\alpha}_{m+1} & \gamma_0 \end{vmatrix} - \gamma_1 \begin{vmatrix} \boldsymbol{\alpha}_0 & x \\ \boldsymbol{\alpha}_2 & \gamma_{m-1} \\ \vdots & \vdots \\ \boldsymbol{\alpha}_{m+1} & \gamma_0 \end{vmatrix} + \cdots
$$

$$
+ (-1)^{1+m+1}\gamma_m \begin{vmatrix} \boldsymbol{\alpha}_0 & x \\ \boldsymbol{\alpha}_1 & \gamma_m \\ \vdots & \vdots \\ \boldsymbol{\alpha}_{m-1} & \gamma_2 \\ \boldsymbol{\alpha}_{m+1} & \gamma_0 \end{vmatrix}
$$

$$
+ (-1)^{1+m+2} x \begin{vmatrix} \boldsymbol{\alpha}_0 & x \\ \boldsymbol{\alpha}_1 & \gamma_m \\ \vdots & \vdots \\ \boldsymbol{\alpha}_m & \gamma_1 \end{vmatrix}
$$

$$
= c_0 - \gamma_1\left[(-1)^{m+2} x \begin{vmatrix} \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_{m-1} \end{vmatrix} + c_1\right] + \cdots
$$

$$+(-1)^{m+2}\gamma_m\left[(-1)^{m+2}x\begin{vmatrix}\alpha_1\\ \vdots\\ \alpha_{m-1}\\ \alpha_{m+1}\end{vmatrix}+c_m\right]$$

$$+(-1)^{m+3}x\begin{vmatrix}\alpha_0 & x\\ \alpha_1 & \gamma_m\\ \vdots & \vdots\\ \alpha_m & \gamma_1\end{vmatrix},$$

则

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}=(-1)^{m+3}\gamma_1\begin{vmatrix}\alpha_2\\ \vdots\\ \alpha_{m+1}\end{vmatrix}+\cdots+(-1)^{2(m+2)}\gamma_m\begin{vmatrix}\alpha_1\\ \vdots\\ \alpha_{m-1}\\ \alpha_{m+1}\end{vmatrix}$$

$$+(-1)^{m+3}\begin{vmatrix}\alpha_0 & x\\ \alpha_1 & \gamma_m\\ \vdots & \vdots\\ \alpha_m & \gamma_1\end{vmatrix}+(-1)^{m+3}x(-1)^{m+2}\begin{vmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_m\end{vmatrix}$$

$$=(-1)^{m+3}\left(\begin{vmatrix}\alpha_0 & x\\ \alpha_1 & \gamma_m\\ \vdots & \vdots\\ \alpha_m & \gamma_1\end{vmatrix}+\begin{vmatrix}\gamma_1 & \alpha_1\\ \vdots & \vdots\\ \gamma_m & \alpha_m\\ x & \alpha_{m+1}\end{vmatrix}\right). \quad \blacksquare$$

推论 若 $\alpha_0=(\gamma_1,\gamma_2,\cdots,\gamma_m)$,

$\alpha_1=(\gamma_0,\gamma_1,\cdots,\gamma_{m-1})$,

$\vdots \qquad \vdots$

$\alpha_{m+1}=(\gamma_m,\gamma_{m-1},\cdots,\gamma_1)$,

则

$$f'(x)=(-1)^{m+3}2\begin{vmatrix}\gamma_1 & \alpha_1\\ \gamma_2 & \alpha_2\\ \vdots & \vdots\\ \gamma_m & \alpha_m\\ x & \alpha_{m-1}\end{vmatrix}.$$

证明 令

$$\boldsymbol{A}=\begin{bmatrix}\alpha_0 & x\\ \alpha_1 & \gamma_m\\ \vdots & \vdots\\ \alpha_m & \gamma_1\end{bmatrix}=\begin{bmatrix}\gamma_1 & \gamma_2 & \cdots & \gamma_m & x\\ \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} & \gamma_m\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \gamma_{m-1} & \gamma_{m-2} & \cdots & \gamma_0 & \gamma_1\end{bmatrix}.$$

则

$$\boldsymbol{A}'=\begin{pmatrix}\gamma_1 & \alpha_1\\ \gamma_2 & \alpha_2\\ \vdots & \vdots\\ \gamma_m & \alpha_m\\ x & \alpha_{m+1}\end{pmatrix}=\begin{pmatrix}\gamma_1 & \gamma_0 & \vdots & \gamma_{m-1}\\ \gamma_2 & \gamma_1 & \vdots & \gamma_{m-2}\\ \vdots & \vdots & \vdots & \vdots\\ \gamma_m & \gamma_{m-1} & \vdots & \gamma_0\\ x & \gamma_m & \vdots & \gamma_1\end{pmatrix}.$$

由于 $\det\boldsymbol{A}=\det\boldsymbol{A}'$，代入引理4.1中即可。 ∎

定理4.6 设给了 $p+1$ 个实数 $\{R_0,R_1,\cdots,R_p\}$，$\boldsymbol{\Gamma}_{p+1}>0$。记

$$\mathscr{K}_p=\{\zeta_t:\ \zeta_t \text{ 非奇异的平稳列，} R_\zeta(k)=R_k, 0\leqslant k\leqslant p\},\tag{4.54}$$

设 $\xi_t\in\mathscr{K}_p$，为了对 $\forall N>p+1$，一切 $\zeta_t\in\mathscr{K}_p$，使得

$$H(\xi_1,\xi_2,\cdots,\xi_N)\geqslant H(\zeta_1,\zeta_2,\cdots,\zeta_N)\tag{4.55}$$

成立的充要条件是：ξ_t 是正态AR(p)型，其方程

$$\sum_{k=0}^{p}\varphi_k\xi_{t-k}=\theta_0\varepsilon_t\tag{4.56}$$

的参数 θ_0 及 $\varphi=(\varphi_1,\varphi_2,\cdots,\varphi_p)'$ 满足由 $\{R_0,R_1,\cdots,R_p\}$ 所组成的 Yule-Walker 方程(4.32)。

证明 先证充分性. 设 ξ_t 是(4.56)的正态AR(p)型. 显然我们只需对 $\mathscr{K}_p$ 中的正态列 ξ_t 检验(4.55)即可. 即对 $N > p+1$, 检验

$$\det \Gamma_N^{(\xi)} \geqslant \det \Gamma_N^{(\zeta)}. \tag{4.57}$$

先看 $N = p+2$, 这时

$$\Gamma_N^{(\xi)} = \begin{vmatrix} R_0 & R_1 & \cdots & R_p & R^{(\xi)}_{p+1} \\ & & & & R_p \\ \vdots & & & \vdots & \vdots \\ R_p & R_{p-1} & \cdots & R_0 & R_1 \\ R^{(\xi)}_{p+1} & R_p & \cdots & R_1 & R_0 \end{vmatrix}. \tag{4.58}$$

问题变为当 $R^{(\xi)}_{p+1}$ 取为 $R^{(\xi)}_{p+1}$ 时上述矩阵行列式的值是极大的. 令

$$f_{p+2}(x) = \begin{vmatrix} R_0 & R_1 & \cdots & R_p & x \\ & & & & R_p \\ \vdots & & \ddots & \vdots & \vdots \\ R_p & \cdots & & R_0 & \\ x & R_p & \cdots & \cdots & R_0 \end{vmatrix} = \begin{vmatrix} R_0 & \boldsymbol{R}_0 & x \\ R_1 & \boldsymbol{R}_1 & R_p \\ \vdots & \vdots & \vdots \\ R_p & \boldsymbol{R}_p & R_1 \\ x & \boldsymbol{R}_{p+1} & R_0 \end{vmatrix}, \tag{4.59}$$

其中 $\boldsymbol{R}_i$ 定义如引理 4.1 推论中的 a_i (只是用 R_i 代替 γ_i). 由引理推论知

$$\frac{\mathrm{d}f_{p+2}(x)}{\mathrm{d}x} = 2(-1)^{p+3} \begin{vmatrix} R_1 & R_0 & R_1 & \cdots & R_{p-1} \\ R_2 & R_1 & R_0 & \cdots & R_{p-2} \\ \vdots & \vdots & \vdots & & \vdots \\ R_p & R_{p-1} & R_{p-2} & \cdots & R_0 \\ x & R_p & R_{p-1} & \cdots & R_1 \end{vmatrix}. \tag{4.60}$$

将(4.60)的各列向量记为

$$\boldsymbol{R}', \ \boldsymbol{R}'_1, \ \boldsymbol{R}'_2, \ \cdots, \ \boldsymbol{R}'_p. \tag{4.61}$$

假定 ξ_t 如定理条件所述是(4.56)所描述的AR(p)模型, 由

$$\Phi(U) R_k^{(\xi)} = 0, \quad k \geqslant 1,$$

$$R_k^{(\xi)}=-\sum_{l=1}^{p}\varphi_l R_{k-l}^{(\xi)},\quad k\geqslant 1。\tag{4.62}$$

对照(4.60)表明：若取 $x=R_{p+1}^{(\xi)}$，恰有

$$\boldsymbol{R}'=-\sum_{l=1}^{p}\varphi_l \boldsymbol{R}'_l。\tag{4.63}$$

此时

$$\left.\frac{\mathrm{d}f_{p+2}(x)}{\mathrm{d}x}\right|_{x=R_{p+1}^{(\xi)}}=0。\tag{4.64}$$

又

$$\begin{aligned}\frac{\mathrm{d}^2 f_{p+2}(x)}{\mathrm{d}x}&=2(-1)^{p+3}\frac{\mathrm{d}}{\mathrm{d}x}\begin{vmatrix} R_1 & \\ R_2 & \boldsymbol{\Gamma}_p \\ \vdots & \\ R_p & \\ x & R_p \cdots R_1\end{vmatrix}\\ &=2(-1)^{p+3}(-1)^{p+2}\det\boldsymbol{\Gamma}_p\\ &=-2\det\boldsymbol{\Gamma}_p<0。\end{aligned}\tag{4.65}$$

表明 $N=p+2$ 时，$\det\boldsymbol{\Gamma}_{p+2}^{(\xi)}$ 达极大值，故

$$H(\xi_1,\xi_2,\cdots,\xi_{p+2})\geqslant H(\zeta_1,\zeta_2,\cdots,\zeta_{p+2})。\tag{4.66}$$

进一步可以证明(参看[84])当 $N>p+L$ $(L>1)$时(4.57)成立。

再证必要性。设 ξ_t 是正态平稳列，满足 $\xi_t\in\mathscr{K}_p$，并对任意的 $N>p+1$，$\det\boldsymbol{\Gamma}_N^{(\xi)}$ 达极大，以下证它必为AR(p)型。

对 $N=p+2$，由条件，应有

$$f_{p+2}^{(\xi)}(x)=\det\boldsymbol{\Gamma}_{p+2}^{(\xi)}。\tag{4.67}$$

因引理4.1中对(4.52)展式知(4.67)是 x 的实系数二次多项式，并且 x^2 项的系数是

$$(-1)\det\boldsymbol{\Gamma}_p<0。\tag{4.68}$$

因此 $f_{p+2}(x)$ 有唯一的极大值 x_0 并 $f'_{p+2}(x_0)=0$，$x_0=R_{p+1}^{(\xi)}$。

由(4.60)得

$$f'_{p+2}(x_0)=2(-1)^{p+3}\left|\begin{pmatrix}R_1\\R_2\\\vdots\\R_p\\x_0\end{pmatrix}\begin{pmatrix}R_0 & R_1 & \cdots & R_{p-1}\\R_1 & R_0 & \cdots & R_{p-2}\\\vdots & \vdots & & \vdots\\R_{p-1} & R_{p-2} & \cdots & R_0\\R_p & R_{p-1} & \cdots & R_1\end{pmatrix}\right|. \qquad (4.69)$$

而 $\det\boldsymbol{\Gamma}_p>0$，所以 $\boldsymbol{R}'_1,\boldsymbol{R}'_2,\cdots,\boldsymbol{R}'_p$ 线性无关。因而不难看出：

$$f'_{p+2}(x_0)=0\Longleftrightarrow-\boldsymbol{R}'=\sum_{l=1}^{p}\varphi_l\boldsymbol{R}'_l. \qquad (4.70)$$

(4.70)由右推左是显然的；又，若 $f'_{p+2}(x_0)=0$，则必有

$$c_0\boldsymbol{R}'+c_1\boldsymbol{R}'_1+\cdots+c_p\boldsymbol{R}'_p=0. \qquad (4.71)$$

但 $c_0\neq0$（因$\boldsymbol{R}'_1,\cdots,\boldsymbol{R}'_p$ 线性无关），故(4.71)可改写为

$$\boldsymbol{R}'=-\sum_{i=1}^{p}\varphi_i\boldsymbol{R}'_i. \qquad (4.72)$$

显然它就是 Yule-Walker 方程。如今 $\det\Gamma^{(\xi)}_{p+2}$ 达极大，则

$$x_0=R^{(\xi)}_{p+1}=-\sum_{l=1}^{p}\varphi_lR_{p+1-l}.$$

对 $N=p+3$，可考查

$$\frac{\mathrm{d}f_{p+3}(x)}{\mathrm{d}x}\overset{①}{=}2(-1)^{p+4}\left|\begin{pmatrix}R_1\\R_2\\\vdots\\R_p\\R_{p+1}\\x\end{pmatrix}\begin{pmatrix}R_0 & \cdots & R_{p-1} & R_p\\R_1 & \cdots & R_{p-2} & R_{p-1}\\\vdots & & \vdots & \vdots\\R_{p-1} & \cdots & R_0 & R_1\\R_p & \cdots & R_1 & R_0\\R_{p+1} & \cdots & R_2 & R_1\end{pmatrix}\right|.$$

(4.73)

然而 $\boldsymbol{\beta}'_1,\cdots,\boldsymbol{\beta}'_{p+1}$($\boldsymbol{\beta}'_i$ 是上式行列式各右列组成的向量)线性无关（因 $\det\boldsymbol{\Gamma}_{p+1}>0$），故有唯一极大值 $x_0=R^{(\xi)}_{p+2}$的充要条件是

① 当足标 $l>p$，简记$R^{(\xi)}_l=R_l$。

$$\boldsymbol{\beta}' = \sum_{l=1}^{p+1} c_l \boldsymbol{\beta}'_l \tag{4.74}$$

或

$$\begin{bmatrix} \boldsymbol{R}' \\ x_0 \end{bmatrix} = c_1 \begin{bmatrix} \boldsymbol{R}'_1 \\ R_{p+1} \end{bmatrix} + c_2 \begin{bmatrix} \boldsymbol{R}'_2 \\ R_p \end{bmatrix} + \cdots + c_p \begin{bmatrix} \boldsymbol{R}'_p \\ R_2 \end{bmatrix} + c_{p+1} \begin{bmatrix} R_p \\ \vdots \\ R_1 \end{bmatrix}.$$

结合(4.72)知

$$\boldsymbol{R}' = \sum_{l=1}^{p+1} c_l \boldsymbol{R}'_l = -\sum_{l=1}^{p} \varphi_l \boldsymbol{R}'_l. \tag{4.75}$$

又

$$(c_1 + \varphi_1)\boldsymbol{R}'_1 + \cdots + (c_p + \varphi_p)\boldsymbol{R}'_p + c_{p+1}\boldsymbol{R}'_{p+1} = 0,$$

由线性无关性只能

$$\begin{cases} c_{p+1} = 0, \\ c_l = -\varphi_l, \quad l = 1, 2, \cdots, p. \end{cases} \tag{4.76}$$

因此

$$x_0 = R^{(\xi)}_{p-2} = -\sum_{l=1}^{p} \varphi_l R_{p-2-l}. \tag{4.77}$$

同理对一般 $N = p + L + 2\ (L \geqslant 1)$，亦可用上法递推。由此，一般说来有

$$\Phi(U) R^{(\xi)}_k = \sum_{l=0}^{p} \varphi_l R^{(\xi)}_{k-l} = 0, \qquad k \geqslant 1.$$

而由第二章的(2.149)—(2.152)的论证知，ξ_t 必是 AR(p)模型

$$\sum_{l=0}^{p} \varphi_l \xi_{t-l} = \theta_0 \varepsilon_t,$$

其中$(\varphi_1, \cdots, \varphi_p)$由(4.72)的 Yule-Walker 方程确定并且

$$\theta_0^2 = \sum_{l=0}^{p} \varphi_l R_l. \quad \blacksquare$$

由定理4.5与定理4.6知，不同的准则却导出相同的结论——

即已知$\{R_k; 0\leqslant k\leqslant p\}$的合理扩充必导致AR模型。但是比较这两条定理还是不尽相同的，联合熵极大的准则显得结果更强一些(对任意的N维分布)，但结论只在正态类中选AR模型，而一步预测误差(或谱熵)极大准则没有限定在正态类中。

§3 实测数据模型拟合和谱估计的判阶问题

1° 实测数据极大熵谱估计中的问题

从定理4.4—定理4.6知，如果已知$R_0, R_1, \cdots, R_p$是x_t非奇异平稳列的$p+1$个相关函数值，则用满足

$$\sum_{l=0}^{p}\varphi_l\xi_{t-l}=\theta_0\varepsilon_t \tag{4.78}$$

的ξ_t来拟合x_t比较合理，因而

$$f_\xi(\lambda)=\frac{\theta_0^2}{2\pi}\Big/\left|\sum_{k=0}^{p}\varphi_k e^{-ik\lambda}\right|^2, \qquad \lambda\in\Pi \tag{4.79}$$

称为极大熵谱估计，$(\theta_0, \varphi_1, \cdots, \varphi_p)$由$\{R_k\}_0^p$确定的Yule-Walker方程解出。

如果对x_t并不预先知道$\{R_k\}_0^p$，而只获得它的实测N个样本值(不间断的测量值)

$$x_1,\ x_2,\ \cdots,\ x_N, \tag{4.80}$$

如何进行(4.79)的估计呢？一种合理的手续是：先求样本相关函数

$$\gamma_k=\frac{1}{N}\sum_{l=1}^{N-k}x_l x_{l+k}, \qquad k=0,1,\cdots,N-1 \tag{4.81}$$

以代替x_t的相关函数$\{R_k\}$。但是当$k=N-1$时(4.81)变成

$$\gamma_{N-1}=\frac{1}{N}x_1x_N, \tag{4.82}$$

它和$R_x(N)=R_N$显然差别太大。因此作为比较合理的近似只能

选 $\gamma_0,\gamma_1,\cdots,\gamma_p$ $(p<N-1)$来替代$\{R_k\}_0^p$，并按定理4.5的手续求$(\theta_0^{(p)},\varphi_1^{(p)},\cdots,\varphi_p^{(p)})$然后用

$$f_p(\lambda)=\frac{(\theta_0^{(p)})^2}{2\pi}\Big/\left|1-\sum_{k=1}^{p}\varphi_k^{(p)}e^{-ik\lambda}\right|^2 \qquad (4.83)$$

来估计 $f_x(\lambda)$。

但是 p 又应该怎样选？这一问题称为定阶问题。

2° 赤池的信息定阶准则(AIC)

关于判阶问题近年来有许多讨论(见[42],[43])，比较广泛流传并且意义比较明确的是日本的赤池(H.Akaike)提出的信息判阶准则 AIC，其基本思想是：

假设 x_t 是 ARMA 模型，参数为θ(含 $p+q+1$ 未知量)，则它的N维分布 $p(x_1,x_2,\cdots,x_N)$ 也应该由θ确定记为 $g(\boldsymbol{x}|\boldsymbol{\theta})$。如果$x_t$的真参数为 $\boldsymbol{\theta}_0$，分布为 $p_0(\boldsymbol{x})=g(\boldsymbol{x}|\boldsymbol{\theta}_0)$，由 Kullback-Leibler 信息量的定理4.1知此时应有

$$I(p_0,g(\boldsymbol{x},\boldsymbol{\theta}_0))=0。\qquad (4.84)$$

如果$\boldsymbol{\theta}_0$是未知的，用统计量$\hat{\boldsymbol{\theta}}$估计它，则应在适当的范围内选$\hat{\boldsymbol{\theta}}$使

$$I(p_0,g(\boldsymbol{x}|\hat{\boldsymbol{\theta}})) \qquad (4.85)$$

愈小愈好。假设 x_t 的真 $\boldsymbol{\theta}_0$ 含有 p 个参数，且由一组样本$(x_1,\cdots,x_N)$出发$\hat{\boldsymbol{\theta}}$ 的形式为已知，参数为 s 个$(p\leqslant s)$，应选一个 $L(N)$ 使满足

$$1\leqslant p\leqslant s\leqslant L(N), \qquad (4.86)$$

则(4.85)变为

$$V(s)=I(p_0,g(\boldsymbol{x}|\hat{\boldsymbol{\theta}}(s))),\quad 1\leqslant s\leqslant L(N)。\qquad (4.87)$$

因而只需选 s 尽可能小并使 $V(s)=\min$。

假定 x_t 是正态 ARMA 模型，上述观点将导致

$$\mathrm{AIC}(s)=\log(\hat{\sigma}_s^2)+2\frac{s}{N},\quad 1\leqslant s\leqslant L(N) \qquad (4.88)$$

达极小值。其中 σ_s^2 为 ARMA 模型的 σ_0^2 或 θ_0^2 的估计值，如果有若干个点达同样的极小值则取最小的 s 值为阶的估计。

具体说来：

ARMA 模型：$s=p+q+1$；

AR 模型：　$s=p+1$。

在给出一组样本$(x_1,x_2,\cdots,x_N)$后如何估计σ_s^2(理论上讲应是极大似然估计)比较复杂，ARMA 模型的场合可参看 Hannan 和 Rissanen 见[44]，[45]。在 AR 模型中由 Yuler-Walker 方程知可取

$$\hat{\sigma}_s^2=\hat{\theta}_{0,S}^2=\hat{\gamma}_0-\sum_{l=1}^{s}\varphi_l^{(s)}\hat{\gamma}_l, \tag{4.89}$$

其中 $\hat{\gamma}_l$ 由(4.81)确定。

继之而来的问题显然是：用(4.88)来判阶能判得准吗？以下介绍这方面的一些结果。

3° AR 模型判阶中的相容性问题

设 x_t 是平稳列并为 AR(p_0)模型，又给了一串 x_t 的不同瞬观测值

$$x_1,\ x_2,\ \cdots,\ x_N.$$

并用(4.81)得 $\hat{\gamma}_0,\hat{\gamma}_1,\cdots,\hat{\gamma}_k$，解Yule-Walker 方程

$$\begin{bmatrix}\hat{\gamma}_0 & \hat{\gamma}_1 & \cdots & \hat{\gamma}_k\\ \vdots & \ddots & \ddots & \vdots\\ & & \ddots & \hat{\gamma}_1\\ \hat{\gamma}_k & \cdots & & \hat{\gamma}_0\end{bmatrix}\begin{bmatrix}1\\ \hat{\varphi}\\ \vdots\\ \hat{\varphi}_k\end{bmatrix}=\begin{bmatrix}\hat{\theta}_0^2(k)\\ 0\\ \vdots\\ 0\end{bmatrix},\qquad 1\leqslant k\leqslant p, \tag{4.90}$$

用 AIC(s)来判阶得 $\hat{p}_N$，当 $N\to\infty$ 时，是否有

$$\hat{p}_N\to p_0 \tag{4.91}$$

成立？上述收敛意义是否是强相容(即 Pr. 1)？至少依概率(弱相容)是否成立？

日本统计学家柴田里程(R.Shibata)(见[46])1976年给出证明，指出了(4.91)连弱相容也不成立，并且一般讲，AIC的定阶可能偏高而不会低。为了介绍他的结果，需要用到二条引理。

引理4.2① 设 x_t 是正态 AR(p)序列，模型参数为

$$(\varphi_1, \varphi_2 \cdots, \varphi_p, \theta_0^2),$$

即 x_t 满足

$$\sum_{i=0}^{p} \varphi_i x_{t-i} = \theta_0 \varepsilon_t,$$

其中 ε_t 为 i.i.d. $N(0,1)$列。对$k>p$，记

$$A_k = (\overbrace{\varphi_1, \varphi_2, \cdots, \varphi_p, 0, \cdots, 0}^{k}),$$

$$\hat{A}_k = (\hat{\varphi}_1, \hat{\varphi}_2, \cdots, \hat{\varphi}_p, \hat{\varphi}_{p+1}, \cdots, \hat{\varphi}_k)\text{(参看(4.90)式)},$$

则

(1) $\hat{A}_k \xrightarrow{a.s} A_k \quad (N \to \infty)$。 (4.92)

(2) $\sqrt{N}(\hat{A}_k - A_k) \xrightarrow{D} N(0, \theta_0^2 \boldsymbol{\Gamma}_k^{-1})$，

其中

$$\boldsymbol{\Gamma}_k = (R(i-j))_{0<i,j<k}。$$

(3) 当 $n>m\geqslant p$ 时，

$$E(\hat{\varphi}_n \hat{\varphi}_m) = o\left(\frac{1}{N}\right)。$$

(4) $N \sum_{k=p+1}^{k+p} (\hat{\varphi}_k)^2$ 的渐近分布为 χ_K^2。 (4.93)

引理4.3(Spitzer F.[47]) 设 $z_1, z_2, \cdots, z_n$ 为i.i.d.的随机变量。令

$$s_k = z_1 + z_2 + \cdots + z_k, \quad 1 \leqslant k \leqslant n, \tag{4.94}$$

则

$$P\left\{\bigcap_{l=1}^{n} (s_l > 0)\right\} = q_n = \sum^{*}\left\{\prod_{i=1}^{n} \frac{1}{r_i!}\left(\frac{\alpha_i}{i}\right)^{r_i}\right\}, \tag{4.95}$$

$$P\left\{\bigcap_{l=1}^{n} (s_l \leqslant 0)\right\} = q'_n = \sum^{*}\left\{\prod_{i=1}^{n} \frac{1}{r_i!}\left(\frac{1-\alpha_i}{i}\right)^{r_i}\right\}, \tag{4.96}$$

---

① 可参看[26]附录。

其中 $a_i=P\{s_i>0\}$，$\sum\limits_{*}^{*}$表示对所有满足$\sum\limits_{l=1}^{n} lr_n=n$的$(r_1,r_2,\cdots,r_n)$求和。

在以上二条引理的基础上，我们可以介绍重要的柴田定理：

定理4.7（柴田，1976） 设 x_t 是正态的满足

$$\sum_{l=0}^{p_0}\varphi_l x_{t-l}=\theta_0\varepsilon_t$$

的 AR(p_0)序列。由样本 $x_1,x_2,\cdots,x_N$ 出发用 AIC 估阶为 $\hat{p}_N$，选阶的范围为 $0\leqslant p\leqslant P$，而 $p_0\leqslant P$，则

$$\begin{cases}\lim\limits_{N\to\infty} P\{\omega:\ \hat{p}_N(\omega)=p\}=q_{p-p_0}\cdot q'_{P-p}, & p_0\leqslant p\leqslant P; \quad (4.97)\\ P\{\omega:\ \lim\limits_{N\to\infty}\hat{p}_N(\omega)=p\}=0, & 0\leqslant p<p_0, \quad (4.98)\end{cases}$$

其中 q_n,q'_n 由(4.95)，(4.96)定义，而

$$\alpha_i=P\{x_i^2>2i\}。\tag{4.99}$$

证明 对 AR 模型讲，利用 Y-W 方程，当 m 不超过阶数 p_0 时，

$$\begin{bmatrix}\theta_0^2(m+1)\\0\\\vdots\\0\\0\end{bmatrix}=\begin{bmatrix}R_0 & R_1 & \cdots & R_{m+1}\\ & R_0 & & \\ \vdots & & \ddots & \vdots\\ R_{m+1} & & \cdots & R_0\end{bmatrix}$$

$$\times\left\{\begin{bmatrix}1\\-\varphi_1^{(m)}\\\vdots\\-\varphi_m^{(m)}\\0\end{bmatrix}-\varphi_{m+1}^{(m+1)}\begin{bmatrix}0\\-\varphi_m^{(m)}\\\vdots\\-\varphi_1^{(m)}\\1\end{bmatrix}\right\}。\tag{4.100}$$

可见（利用(2.121)）

$$\theta_0^2(m+1)=\theta_0^2(m)-\varphi_{m+1}^{(m+1)}\left(R_{m+1}-\sum_{j=1}^{m}\varphi_j^{(m)}R_{m+1-j}\right)$$

$$=\theta_0^2(m)(1-(\varphi_{m+1}^{(m+1)})^2)$$

$$=R_0\prod_{i=1}^{m+1}[1-(\varphi_i^{(i)})^2].\qquad(4.101)$$

如果用$\{\hat{\gamma}_k\}$代替$\{R_k\}$，并把相应参数记为$\{\hat{\varphi}_j^{(i)}\},\hat{\theta}_0$，则

$$\hat{\theta}_0^2(m+1)=\hat{\gamma}_0\prod_{i=1}^{m+1}[1-(\hat{\varphi}_i^{(i)})^2].\qquad(4.102)$$

(1) 取$p\in[0,p_0)$，则由(4.101)得

$$\frac{\hat{\theta}_0^2(p)}{\hat{\theta}_0^2(p_0)}=\prod_{i=p+1}^{p_0}[1-(\hat{\varphi}_i^{(i)})^2]^{-1}$$

$$\xrightarrow{\text{a.s.}}\prod_{i=p+1}^{p_0}[1-(\varphi_i^{(i)})^2]^{-1}\text{(利用引理4.2的1)}.\qquad(4.103)$$

由于矩阵$\boldsymbol{\Gamma}_{m+1}>0$，由(4.101)关系式知，当

$$0<\hat{\theta}_0^2(m+1)=\hat{\theta}_0^2(m)(1-(\hat{\varphi}_{m+1}^{(m+1)})^2),\quad m=0,1,\cdots,k\leqslant p_0$$

时必有$|\hat{\varphi}_{m+1}^{(m+1)}|<1$。再考查(4.103)知

$$c=\prod_{i=p+1}^{p_0}[1-(\varphi_i^{(i)})^2]^{-1}>1.\qquad(4.104)$$

表明在概率空间$(\Omega,\mathscr{F},P)$中存在$\Omega'\subset\Omega$，$P\{\Omega'\}=1$，当$\omega\in\Omega'$时

$$\log\left(\frac{\hat{\theta}^2(p)}{\hat{\theta}_0^2(p_0)}\right)\to\log c>0,\quad\text{a.s.}.\qquad(4.105)$$

假定用AIC法则定阶为$\hat{p}_N$，则对$p\in[0,p_0)$ $\hat{p}_N$选为p意味着如下的随机事件：

$$\{\hat{p}_N=p\}=\left\{\bigcap_{0\leqslant k\leqslant P}(\text{AIC}(p)\leqslant\text{AIC}(k))\right\}\qquad(4.106)$$

$$\subset\{\text{AIC}(p)\leqslant\text{AIC}(p_0)\}\qquad(4.107)$$

$$=\left\{\log(\hat{\theta}_0^2(p))+\frac{2p}{N}\leqslant\log(\hat{\theta}_0^2(p_0))+\frac{2p_0}{N}\right\}$$

$$=\left\{\log\left(\frac{\hat{\theta}_0^2(p)}{\hat{\theta}_0^2(p_0)}\right)\leqslant\frac{2}{N}(p_0-p)\right\}。\tag{4.108}$$

于是

$$\left\{\lim_{N\to\infty}\hat{p}_N=p\right\}\subset\left\{\lim_{N\to\infty}\log\left(\frac{\hat{\theta}_0^2(p)}{\hat{\theta}_0^2(p_0)}\right)\leqslant 0\right\}。\tag{4.109}$$

和(4.105)相比知(4.109)必属于 $\Omega\backslash\Omega'$ 中的随机事件，因而

$$P\{\omega;\ \lim_{N\to\infty}\hat{p}_N(\omega)=p\}=0,\quad p\in[0,p_0)$$

即为(4.98)。

(2) 对 $p\in[p_0,P]$，上面已证 $\hat{p}_N$ 不可能取值在$[0,p_0)$. 因此当N充分大，$N\geqslant N_1$ 后，对 $\varepsilon/2>0$ 有

$$P\left\{\bigcap_{p_0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\}-\frac{\varepsilon}{2}$$

$$\leqslant P\left\{\bigcap_{0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\}=P\{\hat{p}_N=p\}$$

$$\leqslant P\left\{\bigcap_{p_0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\},$$

或

$$\left|P\left\{\bigcap_{0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\}\right.$$

$$\left.-P\left\{\bigcap_{p_0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\}\right|<\frac{\varepsilon}{2}。\tag{4.110}$$

利用(4.103)左端的等式，(4.110)第二项

$$P\left\{\bigcap_{p_0<k<P}(\mathrm{AIC}(p)\leqslant\mathrm{AIC}(k))\right\}$$

$$=P\left\{\bigcap_{p_0<k<P}\left(\log\hat{\theta}_0^2(p)+\frac{2}{N}p\leqslant\log\hat{\theta}_0^2(k)+\frac{2}{N}k\right)\right\}$$

$$= P\left\{\bigcap_{p_0 \leqslant k < p}\left(-N\log\prod_{j=k+1}^{p}(1-(\phi_j^{(j)})^2) > 2(p-k)\right)\right.$$

$$\left.\cdot\bigcap_{p < k \leqslant P}\left(-N\log\prod_{j=p+1}^{k}(1-(\phi_j^{(j)})^2) \leqslant 2(k-p)\right)\right\}.$$

(4.111)

其中注意：(a) 当 $k=p$ 时(4.111)左交集内的事件是必然的，故第二等式第一交集内可少一项。(b) 由 AIC 的规定，当值相等时选低阶的原则在 $p_0 \leqslant k < p = \hat{p}_N$ 条件下只能有 $\text{AIC}(p) < \text{AIC}(k)$。故(4.111)右端第一交集括号内改为不等号。

利用微积分中如下结果：$-1 \leqslant x < 1$ 内，

$$\log(1-x) = -\left(x + \frac{x^2}{2}\frac{1-\theta}{(1-\theta x)^2}\right)\ (0<\theta<1). \quad (4.112)$$

当 x 充分小有

$$0 \leqslant \frac{1-\theta}{(1-\theta x)^2} \leqslant 1.$$

遂得

$$-\log(1-(\phi_j^{(j)})^2) = (\phi_j^{(j)})^2 + \frac{a}{2}(\phi_j^{(j)})^4, \quad 0 \leqslant a \leqslant 1, \quad (4.113)$$

或

$$-N\log(1-(\phi_j^{(j)})^2) = N(\phi_j^{(j)})^2\left[1 + \frac{a}{2}(\phi_j^{(j)})^2\right]. \quad (4.114)$$

利用(4.92)，(4.93)，当 $j > p_0$ 后依分布($N \to \infty$)

$$N(\phi_j^{(j)})^2 \xrightarrow{D} \chi_1^2. \quad (4.115)$$

若记极限分布的随机变量为 Y_j，则它们是 i.i.d.同 χ_1^2 分布，$j = p_0+1, \cdots, P$。将这些结果应用到(4.111)的第一交集得

$$\bigcap_{p_0 \leqslant k < p}\left[\sum_{j=k+1}^{p}(-N\log(1-(\phi_j^{(j)})^2)) > 2(p-k)\right].$$

再令 $l = j-k$ 和 $s = p-k$，则

$$上式=\bigcap_{1\leqslant s\leqslant p-p_0}\left[\sum_{l=1}^{s}(-N\log(1-(\varphi_{p-s+l}^{(p-s+l)})^2)-2)>0\right]$$

$$=\bigcap_{1\leqslant s\leqslant p-p_0}\left[\sum_{l=1}^{s}(-N\log(1-(\varphi_{p-l+1}^{(p-l+1)})^2)-2)>0\right].\tag{4.116}$$

(4.111)第二交集是

$$\bigcap_{p<k\leqslant P}\left[\sum_{j=p+1}^{k}(-N\log(1-(\varphi_j^{(j)})^2)-2)\leqslant 0\right]$$

$$\xlongequal{令 s=k-p}\bigcap_{1\leqslant s\leqslant P-p}\left[\sum_{j=p+1}^{s+p}(-N\log(1-(\varphi_j^{(j)})^2)-2)\leqslant 0\right]$$

$$=\bigcap_{1\leqslant s\leqslant P-p}\left[\sum_{l=1}^{s}(-N\log(1-(\varphi_{l+p}^{(l+p)})^2)-2)\leqslant 0\right].\tag{4.117}$$

利用(4.114),(4.115)，令 $Z_l=Y_{p+1-l}-2$，$l=1,2,\cdots,p-p_0$；$\zeta_l=Y_{l+p}-2,l=1,2,\cdots,P-p$，则(4.116)和(4.117)方括号内事件可用

$$\left[\sum_{l=1}^{s}Z_l>0\right]和\left[\sum_{l=1}^{s}\zeta_l\leqslant 0\right]$$

替代。由于$\{Z_l\}$由$\{Y_{p_0+1},\cdots,Y_p\}$组成，$\{\zeta_l\}$由$\{Y_{p+1},\cdots,Y_P\}$组成，利用$\{Y_j\}$独立性，得

$$P\left\{\bigcap_{1\leqslant s\leqslant p-p_0}\left[\sum_{l=1}^{s}Z_l>0\right]\bigcap_{1\leqslant s\leqslant P-p}\left[\sum_{l=1}^{s}\zeta_l\leqslant 0\right]\right\}$$

$$=P\left\{\bigcap_{1\leqslant s\leqslant p-p_0}[U_s>0]\bigcap_{1\leqslant s\leqslant P-p}[V_s\leqslant 0]\right\}$$

$$=P\left\{\bigcap_{1\leqslant s\leqslant p-p_0}(U_s>0)\right\}P\left\{\bigcap_{1\leqslant s\leqslant P-p}(V_s\leqslant 0)\right\}$$

$$= q_{p-p_0} \cdot q'_{P-p}, \tag{4.118}$$

其中

$$U_s = \sum_{l=1}^{s} Z_l = \sum_{l=1}^{s} Y_{p+1-l} - 2s, \tag{4.119}$$

$$V_s = \sum_{l=1}^{s} \zeta_l = \sum_{l=1}^{s} Y_{l+p} - 2s. \tag{4.120}$$

(4.119)和(4.120)求和的变量遵从 χ_s^2 分布。故(4.118)的 q, q' 用(4.95)和(4.96)表达时

$$\alpha_i = P\{\chi_i^2 > 2i\}.$$

利用以上结果，当N充分大，$N \geqslant N_2$，对$\frac{\varepsilon}{2}$可有

$$\left| P\left\{ \bigcap_{p_0 < k < P} (\mathrm{AIC}(p) \leqslant \mathrm{AIC}(k)) \right\} - q_{p-p_0} q'_{P-p} \right| < \frac{\varepsilon}{2}. \tag{4.121}$$

所以当 $N \geqslant \max(N_1, N_2)$时，

$$|P\{\hat{p}_N = p\} - q_{p-p_0} q'_{P-p}| < \frac{\varepsilon}{2}, \quad p_0 \leqslant p \leqslant P,$$

则

$$\lim_{N \to \infty} P\{\hat{p}_N = p\} = q_{p-p_0}, \ q'_{P-p}, \quad p_0 \leqslant p \leqslant P. \ \blacksquare$$

从柴田定理结果看 $\hat{p}_N$ 的极限分布与AR(p)模型的参数无关而只取决于 p_0, P。

柴田虽然证明了AIC估计并非真阶的相容估计，但在应用中大量事实表明它还是很有效的判阶方法，以下的实际计算表明在常用范围内，非相容的概率并不大：设真阶 $p_0 \in [0, 10]$，AIC(s)，在$0 \leqslant s \leqslant P$，$P = 10$，内选阶，则

$$\lim_{N \to \infty} P\{\hat{p}_N = p\}$$

诸概率值如表4.1。

表 4.1

p_0 \ p	0	1	2	3	4	5	6	7	8	9	10
0	.7171	.1131	.0577	.0351	.0232	.0162	.0117	.0087	.0067	.0054	.0049
1	0	.7188	.1134	.0580	.0353	.0234	.0164	.0120	.0090	.0072	.0064
2	0	0	.7210	.1139	.0583	.0356	.0238	.0167	.0124	.0097	.0085
3	0	0	0	.7241	.1146	.0588	.0361	.0242	.0173	.0132	.0115
4	0	0	0	0	.7285	.1156	.0596	.0369	.0251	.0186	.0157
5	0	0	0	0	0	.7349	.1171	.0608	.0382	.0269	.0220
6	0	0	0	0	0	0	.7446	.1196	.0630	.0409	.0319
7	0	0	0	0	0	0	0	.7602	.1239	.0674	.0485
8	0	0	0	0	0	0	0	0	.7874	.1326	.0800
9	0	0	0	0	0	0	0	0	0	.8427	.1573
10	0	0	0	0	0	0	0	0	0	0	1

由于柴田的证明引起了许多统计学家对 AIC 的修改。赤池提出

$$\mathrm{BIC}(s)=\log(\hat{\sigma}_0^2(s))+\frac{s}{N}\log N. \qquad (4.122)$$

安，陈和 Hannan 在[43]中证明了 BIC 是相容估计。

定理4.8(An-Chen-Hannan) 设 x_t 是正态① 的AR(p_0)序列，$x_1,\cdots,x_N$ 为样本。令

$$P(N)=O(\log N). \qquad (4.123)$$

则在$[0,P(N)]$范围内运用 BIC(s) 极小化得到的阶 $\hat{p}_N$ 满足

$$\hat{p}_N\to p_0, \quad \text{a.s.}. \qquad (4.124)$$

Hannan 和 Quinn 在[48]中证明，

$$\mathrm{HIC}(s)=\log(\hat{\sigma}_0^2(s))+s\frac{\log\log N}{N}c, \quad c>2 \quad (4.125)$$

在正态① AR 条件下的判阶 $\hat{p}_N$ 也是强相容估计。

虽然 BIC，HIC 理论上解决了估计的强相容性，但在实用中给

① 原定理并不限于正态序列，这里是一种简化叙述。

定样本N之后，研究应该在什么范围求极小值仍沒有解决，因为，以(4.123)为例，$O(\log N)$只是一种对$P(N)$速度增长的控制，并不能给出具体的数值。

4° 长阶自迴归的模型拟合和谱估计

由于判阶的问题使人们考虑：由于现实世界中对x_t的认识本来就不充分，很可能它本身就沒有模型，因此实际的需要是给x_t一种拟合。谱分析的需要就是对$f_x(\lambda)$的一种估计。由此产生以下的想法：即寻求长阶自迴归$\mathrm{AR}(p_N)$，$p_N\to\infty$，$N\to\infty$，使它的谱充分逼近$f_x(\lambda)$。

设x_t是严平稳列并满足一些适当的条件，Berk 在[49]中指出，由$x_1,x_2,\cdots,x_N$样本出发，假定$\hat\varphi_j^{(k)},j=1,2,\cdots,k$，$\hat\theta_0^2(k)$是由$\hat\gamma_0,\hat\gamma_1,\cdots,\hat\gamma_k$确定的 Y-W 方程的解，令

$$\hat f_N(\lambda)=\frac{1}{2\pi}\frac{(\hat\theta_0(p_N))^2}{\left|1-\sum_{k=1}^{p_N}\hat\varphi_k^{(p_N)}\mathrm{e}^{-ik\lambda}\right|^2},\qquad(4.126)$$

其中$p_N=o(\sqrt[3]{N})$，则

$$\hat f_N(\lambda)\xrightarrow{\mathrm{P_r}}f_x(\lambda)\qquad(4.127)$$

安，陈和 Hannan 在[43]中给出了强收敛的长阶自迴归谱估计。正态条件下可叙述为

定理4.9(An–Chen–Hannan) 设x_t是正态平稳列，其 Wold 分解为

$$x_t=\sum_{k=0}^{\infty}c_k\varepsilon_{t-k},\qquad(4.128)$$

满足：$\{c_k\}\in l_1$并且

$$\sum_{k=0}^{\infty}c_kz^k\neq0,\quad|z|\leqslant1.\qquad(4.129)$$

选正整数$p_N\nearrow+\infty$。对某个$\delta>0$，

$$p_N = O(N^{\frac{1}{3}}/\{\ln N(\ln\ln N)^{1+\delta}\}), \tag{4.130}$$

则

$$\sup_{\lambda} |\hat{f}_N(\lambda) - f_x(\lambda)| = o(1), \tag{4.131}$$

其中 $\hat{f}_N(\lambda)$ 由(4.126)所定义。

在应用中，(4.130)的 p_N 可简化为

$$p_N = O(N^{0.3}) = KN^{0.3}, \tag{4.132}$$

而 $2 \leqslant K \leqslant 5$ 的范围是可考虑的区间。

§4 MA模型拟合和谱估计

在§2，§3中主要介绍了AR模型的拟合和谱估计问题。但是有些研究对象 x_t，如果它的特性偏离AR型比较远而样本量不是很大的情况下，用极大熵或长阶自迴归拟合及谱估计效果都不很理想，例如Priestly在[50]中给出了一个例子

$$x_t = \varepsilon_t + 0.95\varepsilon_{t-1}, \quad t = 1, 2, \cdots, 500, \tag{4.133}$$

ε_t 是 i.i.d. $N(0,1)$。用极大熵的AR谱估计法，AIC判阶为 $\hat{p} = 16$，得到的谱估计的谱密度如图4.1所示。

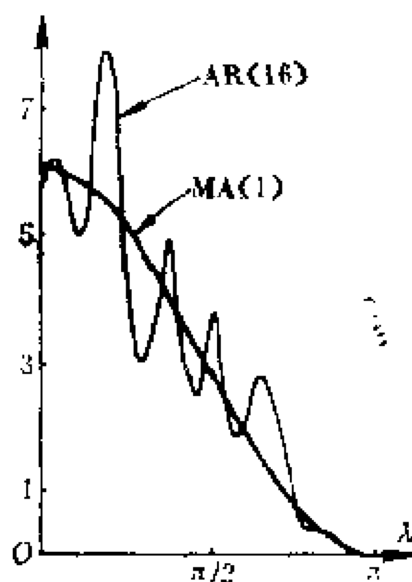


图 4.1

由上看出 AR(16)对 MA(1)谱的估计效果是很差的，因此有些场合就放弃 AR 而考虑 MA(q)模型的拟合和谱估计。

1° 直接法的 MA 拟合和谱估计

设 x_t 为 MA(q)模型，q 已知，则由 $x_1, x_2, \cdots, x_N$ 一组观测样本出发可用第二章和前几节介绍的理论和方法给出 x_t 的拟合模型和谱估计。

(a) 由样本 $\{x_i\}_1^N$ 用(4.81)式计算 x_t 的样本相关函数值 $\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_q$。

(b) 利用相关函数的性质(2.93)式，解如下的关于 $\{\theta_l\}_0^{q+1}$ 的二次方程组

$$\hat{\gamma}_n = \sum_{k=n}^{q} \theta_k \theta_{k-n}, \qquad 0 \leqslant n \leqslant q, \tag{4.134}$$

即

$$\begin{cases} \theta_0^2 + \theta_1^2 + \cdots + \theta_q^2 = \hat{\gamma}_0, \\ \theta_n \theta_0 + \theta_{n+1}\theta_1 + \cdots + \theta_q \theta_{q-n} = \hat{\gamma}_n, \\ \theta_q \theta_0 = \hat{\gamma}_q. \end{cases} \tag{4.135}$$

不难看出，由于 $\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_q, \cdots, \hat{\gamma}_{N-1}$ 是非负定序列，则上述二次方程组必有解。具体算法上可采用如下的叠代解法：

记 $\boldsymbol{\theta}(0) = (\overbrace{0, 0, \cdots, 0}^{q+1})$ 为叠代初值(分量记为$\{\theta_l(0)\}_1^{q+1}$)。

利用(4.135)式，第 $k+1$ 次叠代

$$\boldsymbol{\theta}(k+1) = (\theta_0(k+1), \cdots \theta_q(k+1))$$

是

$$\begin{cases} \theta_0(k+1) = \left(\hat{\gamma}_0 - \sum\limits_{\mu=1}^{q} \theta_\mu^2(k)\right)^{\frac{1}{2}}, \\ \theta_l(k+1) = \left(\hat{\gamma}_l - \sum\limits_{\mu=l+1}^{q} \theta_\mu(k)\theta_{\mu-l}(k)\right) / \theta_0(k+1), \\ \qquad l = 1, 2, \cdots, q-1, \\ \theta_q(k+1) = \hat{\gamma}_q / \theta_0(k+1). \end{cases} \tag{4.136}$$

对给定的叠代精度 $\varepsilon>0$，令

$$\delta(k+1)=\sum_{l=0}^{q}\left|\sum_{\mu=l}^{q}\theta_{\mu}(k+1)\theta_{\mu-l}(k+1)-\hat{\gamma}_l\right|. \tag{4.137}$$

当 $\delta(k+1)<\varepsilon$ 时，$\boldsymbol{\theta}(k+1)$即为诸模型参数的近似解。

在[26]中介绍了用 Newton–Raphson 算法解非线性方程组(4.135)，其收敛速度比(4.136)要快。

(c) 用本章末所附录的判根方法检验(4.135)的解 $\theta_0,\theta_1,\cdots,\theta_q$所组成的多项式

$$\Theta(z)=\sum_{l=0}^{q}\theta_l z^k$$

是否在单位圆内无根。

如果根不在单位圆内，

$$\sum_{l=0}^{q}\theta_l\varepsilon_{t-l}=\hat{x}_t \tag{4.138}$$

就是 x_t 的广 MA(q)拟合模型。

$$\hat{f}_N(\lambda)=\frac{1}{2\pi}\left|\sum_{l=0}^{q}\theta_l e^{-il\lambda}\right|^2,\quad \lambda\in\Pi \tag{4.139}$$

即为 x_t 的 MA 谱估计。[37]指出 $\hat{f}_N(\lambda)$可以是 $f_x(\lambda)$的相容估计。

在应用中，如果 x_t 是 MA(q)但阶数 q 是未知的场合下，可用以下二种方法定阶：

(a) Box–Jenkins 检验。由第二章的习题 9 知：若 x_t 为正态 MA(q)型，则估计 $\hat{\rho}_k$的方差① 是

$$D(\hat{\rho}_k)=\begin{cases}\dfrac{2}{N}\left(\gamma_0^2+2\sum_{l=1}^{q}\gamma_l^2\right)+O\left(\dfrac{1}{N^2}\right), & k=0;\\ \dfrac{1}{N}\left(\gamma_0^2+2\sum_{l=1}^{q}\gamma_l^2\right)+O\left(\dfrac{1}{N^2}\right), & |k|>q.\end{cases} \tag{4.140}$$

① 关于样本协方差的统计性质可参考[77]。

特别是 $k>q$ 以后渐近地有

$$\hat{\rho}_k \sim N\left(0, \frac{1}{N}\left(\gamma_0^2 + 2\sum_{l=1}^{q}\gamma_l^2\right)\right), \quad k>q. \qquad (4.141)$$

参见[26]。因此可先据$\{\hat{\rho}_k\}$的变化估计一个可能的 q 值而后作如下的检验：

H_0: $R_{q+k}=0$ (对 $k=0,1,2,\cdots,[\sqrt{N}]$逐个检验)。

由(4.141)，对显著性水平 $\alpha=0.05$，当 $k=0$时若 $\hat{\rho}_q \in R_\alpha$，即

$$|\hat{\rho}_q| > 2\sqrt{\frac{1}{N}\left(\hat{\rho}_0^2 + 2\sum_{l=1}^{q-1}\hat{\rho}_l^2\right)} \quad (\text{否定}H_0).$$

而 $\{\hat{\rho}_{q+k}, k=1,2,\cdots,[\sqrt{N}]\}$中有不少于$(1-\alpha)\sqrt{N}$个满足

$$|\hat{\rho}_{q+k}| \leqslant 2\sqrt{\frac{1}{N}\left(\hat{\rho}_0^2 + 2\sum_{l=1}^{q}\hat{\rho}_l^2\right)} \qquad (4.142)$$

的即可认为 x_t 是 q 阶相关函数截尾，即阶数为 q 。

(b) 用 AIC 来定阶：由(4.88)，则可选

$$\mathrm{AIC}(s) = \log(\hat{\sigma}_s^2) + 2\frac{s}{N}, \quad 1 \leqslant s \leqslant L(N), \qquad (4.88)$$

$s=q+1$个参数。但在极小化(4.88)时$2\frac{s}{N}=\frac{2q}{N}+\frac{2}{N}$，后一项不变化。故仍取

$$\mathrm{AIC}(q) = \log(\hat{\sigma}_q^2) + 2\frac{q}{N}, \quad 0 \leqslant q \leqslant L(N). \qquad (4.143)$$

而实用的一种近似是取 $\hat{\sigma}_q^2$ 为 MA(q)的一步预测误差，即

$$\hat{\sigma}_q^2 \simeq (\hat{\theta}_0)^2 \qquad (4.144)$$

2° 反相关函数法的 MA 拟合和谱估计

以上介绍的直接法是比较常用的拟合方法，比较直观，而且用(4.139) 作谱估计效果也不错。但最大的一个缺点是：由非线

性方程组(4.135)解出的一组根$\{\theta_0,\theta_1,\cdots,\theta_q\}$未必能保证

$$\hat{\Theta}_N(z)=\sum_{l=0}^{q}\hat{\theta}_l z^l \neq 0,\quad |z|\leqslant 1. \tag{4.145}$$

虽然此时作为谱估计(4.139)仍然是相容估计(见[37])，样本充分大之下$\hat{\Theta}_N(z)$在$|z|=1$上可做到不为零，但在$|z|<1$内仍可能有根，因而(4.138)就不能称为是MA(q)模型。

为解决这一问题Cleveland 1972(见[51])提出了反相关函数，运用这一手法就可以将MA看为AR来处理，而我们知道：运用Yule-Walker方程解出的参数组成的多项式，根都在单位圆外(见(4.40)~(4.40)′)。

定理4.10 设x_t为MA(q)模型

$$\sum_{k=0}^{q}\theta_k\varepsilon_{t-k}=x_t \tag{4.146}$$

的平稳列，

$$\Theta(z)=\sum_{k=0}^{q}\theta_k z^k\neq 0,\qquad |z|\leqslant 1.$$

令①

$$R_i(k)=\int_{\Pi}e^{ik\lambda}\frac{d\lambda}{4\pi^2 f_x(\lambda)},\quad k=0,1,2,\cdots, \tag{4.147}$$

则有方程

$$\begin{bmatrix} R_i(0) & R_i(1) & \cdots & R_i(q)\\ R_i(1) & R_i(0) & \cdots & R_i(q-1)\\ \vdots & & \ddots & \vdots\\ R_i(q) & R_i(q-1) & \cdots & R_i(0)\end{bmatrix}\begin{bmatrix}\theta_0\\ \theta_1\\ \vdots\\ \theta_q\end{bmatrix}=\begin{bmatrix}\frac{1}{\theta_0}\\ 0\\ \vdots\\ 0\end{bmatrix} \tag{4.148}$$

成立。

---

① 通常称$\{R_i(k)\}$为$f_x(\lambda)$或x_t的反相关函数。

证明 由 MA 谱的表达式可得

$$\frac{1}{4\pi^2 f_x(\lambda)} = \frac{1}{2\pi} \frac{1}{\left|\sum_{k=0}^{q}\theta_k e^{-ik\lambda}\right|^2}$$

$$= \frac{1}{2\pi} \frac{\left(\frac{1}{\theta_0}\right)^2}{\left|\sum_{k=0}^{q}\tilde{\theta}_k e^{-ik\lambda}\right|^2}, \tag{4.149}$$

其中

$$\tilde{\theta}_k = \theta_k / \theta_0, \quad k = 0, 1, 2, \cdots, q.$$

$1/4\pi^2 f_x(\lambda)$ 是某个 AR(q)序列的谱，记该序列为 y_t。由(4.149)看出即 y_t 满足如下的AR(q)方程：

$$\sum_{k=0}^{q}\tilde{\theta}_k y_{t-k} = \frac{1}{\theta_0}\varepsilon_t, \tag{4.150}$$

并且

$$R_y(\tau) = R_i(\tau), \quad \tau = 0, \pm 1, \cdots. \tag{4.151}$$

$\{R_y(\tau)\}$满足如下的 Y-W 方程

$$\begin{bmatrix} R_i(0) & R_i(1) & \cdots & R_i(q) \\ \vdots & R_i(0) & & \vdots \\ & & \ddots & \\ R_i(q) & \cdots & \cdots & R_i(0) \end{bmatrix} \begin{bmatrix} 1 \\ \tilde{\theta}_1 \\ \vdots \\ \tilde{\theta}_q \end{bmatrix} = \begin{bmatrix} \frac{1}{\theta_0^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.152}$$

两边乘以 θ_0 即得(4.148)。

推论 设 x_t 为 AR(p)，则 $R_i(\tau)$满足

$$R_i(\tau) = 0, \quad |\tau| > p, \tag{4.153}$$

并且

$$R_i(k) = \frac{1}{\theta_0^2}\sum_{j=0}^{p-k}\varphi_j\varphi_{j+k}, \quad k = 0, 1, \cdots, p, \tag{4.154}$$

其中$\{\theta_0^2, \varphi_1, \cdots, \varphi_p\}$为$x_t$的AR($p$)模型参数。

证明 因为$R_i(k)$对应的谱密度为

$$f_y(\lambda) = \frac{1}{4\pi^2 f_x(\lambda)} = \frac{1}{2\pi}\left|\frac{1}{\theta_0}\sum_{k=0}^{p}\varphi_k e^{-ik\lambda}\right|^2$$

$$= \frac{1}{2\pi}\left|\sum_{k=0}^{p}\left(\frac{\varphi_k}{\theta_0}\right)e^{-ik\lambda}\right|^2, \tag{4.155}$$

因而可表

$$y_t = \sum_{k=0}^{p}\left(\frac{\varphi_k}{\theta_0}\right)\varepsilon_{t-k} \tag{4.156}$$

为MA(p)序列，$\{\varepsilon_t\}$是标准白噪声序列。从而

$$R_y(\tau) = R_i(\tau) = \sum_{k=0}^{p-\tau}\left(\frac{\varphi_k}{\theta_0}\right)\left(\frac{\varphi_{k+\tau}}{\theta_0}\right)$$

$$= \frac{1}{\theta_0^2}\sum_{k=0}^{p-\tau}\varphi_k\varphi_{k+\tau}, \quad \tau = 0, 1, 2, \cdots, p, \tag{4.157}$$

$$R_y(\tau) = R_i(\tau) = 0, \quad 当|\tau| > p. \quad \blacksquare$$

由以上结果知：对MA的拟合，关键是如何得到x_t的反相关函数$\{R_i(\tau)\}$。如果有，只需解(4.152)的Y-W方程即可，从而保证了$\Theta(z)$根在单位圆外的性质。

由一组样本$x_1, x_2, \cdots, x_N$出发，如何运用反相关函数法估计参数$\{\hat{\theta}_k\}$并要求$\hat{\Theta}(z) \neq 0, |z| < 1$，可由以下两种方法进行：

(a) 长阶自回归方法：此法见于Priestly的书[50]。具体方法步骤可采用：

(i) 运用定理4.9，选$p_N = KN^{0.3}$，对$\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_{p_N}$解Y-W方程得参数

$$(\hat{\theta}_0^2(p_N), \hat{\varphi}_1, \cdots, \hat{\varphi}_{p_N}), \tag{4.158}$$

并可认为x_t为AR(p_N)所拟合。

(ii) 计算$\hat{R}_i(\tau)$。对具有(4.158)的AR(p_N)模型运用定理4.10的系，则

$$
\hat{R}_i(k)=\begin{cases}\dfrac{1}{\hat{\theta}_0^2(p_N)}\sum_{j=0}^{p_N-k}\hat{\varphi}_j\hat{\varphi}_{j+k}, & k=0,1,\cdots,p_N;\\ 0, & 其它。\end{cases} \tag{4.159}
$$

(iii) 解 Yule-Walker 方程(定理4.10)

$$
\begin{bmatrix} \hat{R}_i(0) & \hat{R}_i(1) & \cdots & \hat{R}_i(q) \\ \vdots & \ddots & & \vdots \\ \hat{R}_i(q) & \cdots & & \hat{R}_i(0) \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix} = \begin{bmatrix} \frac{1}{\theta_0} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.160}
$$

由此解出的$(\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_q)$即为原来 x_t 的 MA(q)模型的参数估计，x_t的拟合模型为

$$
\tilde{x}_t=\sum_{l=0}^{q}\hat{\theta}_l\varepsilon_{t-l}. \tag{4.161}
$$

谱估计为

$$
f_{\tilde{x}}(\lambda)=\frac{1}{2\pi}\left|\sum_{l=0}^{q}\hat{\theta}_l e^{-il\lambda}\right|^2. \tag{4.162}
$$

如果 x_t 的 MA 阶未知，则一种合理的估计是在(4.160)中看 q 为$[0,L(N)]$内的参变量而运用 AIC 准则来定阶。

(b) 用快速 Fourier 变换(FFT)计算反相关函数法，此法见[52]。

由(4.135)的解可得一组参数估计 $\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_q$，虽然$\hat{\theta}(z)=\sum_{l=0}^{q}\hat{\theta}_l z^l$ 不能保证根皆在单位圆外，但在N比较大时可使

$$
\hat{f}_N(\lambda)=\frac{1}{2\pi}\left|\sum_{l=0}^{q}\hat{\theta}_l e^{-il\lambda}\right|^2>0,\quad \lambda\in\Pi, \tag{4.163}
$$

且可改写为

$$\hat{f}_N(\lambda) = \frac{1}{2\pi}\left|\sum_{k=0}^{q}\hat{\theta}_k^* e^{-ik\lambda}\right|^2, \tag{4.164}$$

$$\sum_{k=0}^{q}\hat{\theta}_k^* z^k \neq 0, \quad |z| \leqslant 1, \tag{4.165}$$

从而表明(4.163)是 MA(q)模型的谱密度。

事实上，设

$$\hat{\Theta}(z) = \hat{\theta}_0(z-\alpha_1)\cdots(z-\alpha_m)(z-\alpha_{m+1})\cdots(z-\alpha_q), \tag{4.166}$$

其中

$$|\alpha_i| < 1, \quad i = 1,2,\cdots,m; \quad |\alpha_j| > 1, \quad j = m+1,\cdots,q. \tag{4.167}$$

则

$$\begin{aligned}
|\hat{\Theta}(e^{-i\lambda})|^2 &= |\hat{\theta}_0|^2|e^{-i\lambda}-\alpha_1|^2\cdots|e^{-i\lambda}-\alpha_m|^2 \\
&\quad \times |e^{-i\lambda}-\alpha_{m+1}|^2\cdots|e^{-i\lambda}-\alpha_q|^2 \\
&= |\hat{\theta}_0|^2\prod_{l=1}^{m}|\alpha_l|^2\left|\frac{1}{\alpha_l}-e^{i\lambda}\right|^2\prod_{k=m+1}^{q}|\alpha_k-e^{-i\lambda}|^2 \\
&= \beta_0^2\prod_{l=1}^{m}\left|\frac{1}{\bar{\alpha}_l}-e^{-i\lambda}\right|^2\prod_{k=m+1}^{q}|\alpha_k-e^{-i\lambda}|^2 \\
&= \beta_0^2\prod_{l=1}^{q}|\beta_l-e^{-i\lambda}|^2,
\end{aligned} \tag{4.168}$$

其中

$$\beta_0^2 = |\hat{\theta}_0|^2\prod_{l=1}^{m}|\alpha_l|^2, \tag{4.169}$$

$$\beta_l = \begin{cases}(\bar{\alpha}_l)^{-1}, & \text{当}|\alpha_l|<1; \\ \alpha_l. & \text{当}|\alpha_l|>1.\end{cases} \tag{4.170}$$

故

$$\hat{\Theta}^*(z) = \beta_0\prod_{l=1}^{q}(\beta_l - z) \neq 0, \quad |z| \leqslant 1, \tag{4.171}$$

即，

$$\hat{f}_N(\lambda)=\frac{1}{2\pi}|\hat{\Theta}^*(e^{-i\lambda})|^2,\quad \lambda\in\Pi. \tag{4.172}$$

显见，运用(4.163)很容易由$(\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_q)$计算出

$$\hat{f}_N(\lambda_k),\quad k=1,2,\cdots,2^M.$$

由此可算出反谱

$$g(\lambda_k)=\frac{1}{4\pi^2}\hat{f}_N(\lambda_k),\quad k=1,2,\cdots,2^M. \tag{4.173}$$

对$\{g(\lambda_k)\}$用 FFT 很快可求出反相关函数$\hat{R}_i(\tau)$值(见[53])。然后再照(4.160)和(4.162)计算即可。

§5 ARMA模型的拟合和谱估计

ARMA 模型的识别和模型拟合不论在理论上还是在应用上都是一个十分重要的问题。但由于出现了 AR 和 MA 两部分，因而使问题变得复杂和困难。目前这一课题的探讨虽已有不少文章，如 Hannan 和 Rissanen(见[54])，王和陈(见[55])，但无论在理论上还是在方法上都不是令人满意的，因而它仍然是近代的一大研究课题。以下只是介绍常用的方法。

1° ARMA 模型参数估计和谱估计的线性方法①

设x_t是 ARMA(p,q)模型，p,q为已知。由$x_1,x_2,\cdots,x_N$样本出发，可按以下步骤进行谱估计：

(a) 由$x_1,\cdots,x_N$计算样本相关函数$\hat{\rho}_0,\hat{\rho}_1,\cdots,\hat{\rho}_{p+q}$。

(b) 由第二章(2.91)式，可用

$$\sum_{k=0}^{p}\varphi_k\hat{\rho}_{n-k}=0,\quad n=q+1,q+2,\cdots,q+p, \tag{4.174}$$

即通常所称的延伸 Yule-Walker 方程组解出$(\varphi_1,\varphi_2,\cdots,\varphi_p)$(因

① 可参看[37]。

为这里假定 p 是已知的)。

(c) 令 $y_t=\sum_{k=0}^{p}\varphi_k x_{t-k}$。对照原来的 ARMA 模型

$$\sum_{k=0}^{p}\varphi_k x_{t-k}=\sum_{l=0}^{q}\theta_l\varepsilon_{t-l}$$

知 y_t 应为 MA(q) 模型。若记 $R_y(\tau)$ 为 y_t 的相关函数，则 $\{R_y(\tau)\}$ 应满足(4.135)类型的如下二次方程组

$$R_y(n)=\sum_{k=n}^{q}\theta_k\theta_{k-n},\quad 0\leqslant n\leqslant q.$$

若已得 $\{\hat{\varphi}_1,\hat{\varphi}_2,\cdots,\hat{\varphi}_p\}$ 诸参数的估计，则用

$$\hat{\gamma}_y(n)=\sum_{k=0}^{p}\sum_{l=0}^{p}\hat{\varphi}_k\hat{\varphi}_l\hat{\gamma}_{n-k+l},\quad \hat{\varphi}_0=1,\quad n=0,1,2,\cdots,q \tag{4.175}$$

来估计 $R_y(n)$ 是合理的，于是用方程组

$$\hat{\gamma}_y(n)=\sum_{k=n}^{q}\hat{\theta}_k\hat{\theta}_{k-n},\quad n=0,1,2,\cdots,q \tag{4.176}$$

的解 $(\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_n)$ 作为估计。

(d) x_t 谱密度的估计为

$$\hat{f}_N(\lambda)=\frac{1}{2\pi}\left|\frac{\sum_{l=0}^{q}\hat{\theta}_l e^{-il\lambda}}{\sum_{k=0}^{p}\hat{\varphi}_k e^{-ik\lambda}}\right|^2. \tag{4.177}$$

在类似正态等条件下，$\hat{\gamma}_k\to\gamma_k$ 是相容估计，因而 $(\hat{\varphi}_1,\hat{\varphi}_2,\cdots,\hat{\varphi}_p)$ 和 $(\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_q)$ 均为 $\{\varphi_k\}_1^p$ 和 $\{\theta_l\}_0^q$ 的相容估计(见[37])，因而 $\hat{f}_N(\lambda)$ 也是。

如果 $(\hat{\theta}_0,\hat{\theta}_1,\cdots,\hat{\theta}_n)$ 能保证

$$\hat{\theta}(z)=\sum_{k=0}^{q}\hat{\theta}_k z^k\neq 0,\quad |z|\leqslant 1, \tag{4.178}$$

则对 x_t 的拟合模型是满足方程

$$\sum_{k=0}^{p}\hat{\varphi}_k\tilde{x}_{t-k}=\sum_{l=0}^{q}\hat{\theta}_l\varepsilon_{t-l} \tag{4.179}$$

的 ARMA(p,q) 序列。

(e) 如果 x_t 的阶 p,q 未知，则可用 AIC 或 BIC 来判阶

$$\mathrm{AIC}(p,q)=\log(\hat{\theta}_0^2(p,q))+2\frac{p+q}{N},\quad 0\leqslant p\leqslant P,0\leqslant q\leqslant Q. \tag{4.180}$$

P,Q 为 ARMA 模型的可能的上界。(4.180) 应逐个地选(p,q)，而 $\hat{\theta}_0^2(p,q)$为相应于(4.176)的第一个参数。若(p_0,q_0)使

$$\mathrm{AIC}(p_0,q_0)=\min_{\substack{0\leqslant p\leqslant P\\0\leqslant q\leqslant Q}}\{\mathrm{AIC}(p,q)\}, \tag{4.181}$$

则定拟合的 ARMA 模型的阶数为(p_0,q_0)。

2° ARMA 模型的谱估计

如果只是为对 ARMA(p,q)的 x_t 进行谱估计，则可运用另一简单方法(p,q假定是已知的)。

定理4.11 设 x_t 为满足 ARMA 方程

$$\sum_{k=0}^{p}\varphi_k x_{t-k}=\sum_{l=0}^{q}\theta_l\varepsilon_{t-l} \tag{4.182}$$

的序列，令

$$y_t=\sum_{l=0}^{q}\theta_l\varepsilon_{t-l}, \tag{4.183}$$

其相关函数记为 $R_y(0),R_y(1),\cdots$,则

$$f_x(\lambda)=\frac{R_y(0)+2\sum_{k=1}^{q}R_y(k)\cos k\lambda}{2\pi\left|\sum_{k=0}^{p}\varphi_k e^{-ik\lambda}\right|^2}. \tag{4.184}$$

证明 由(4.183)知，y_t 的谱密度应是

$$\frac{1}{2\pi}\left|\sum_{l=0}^{q}\theta_l e^{-il\lambda}\right|^2=\frac{1}{2\pi}\sum_{k=-q}^{q}R_y(k)e^{-ik\lambda}.$$

$$\left|\sum_{l=0}^{q}\theta_l e^{-il\lambda}\right|^2=R_y(0)+\sum_{k=1}^{q}R_y(k)(e^{-ik\lambda}+e^{ik\lambda})$$

$$= R_y(0) + 2\sum_{k=1}^{q} R_y(k)\cos k\lambda. \qquad (4.185)$$

而

$$f_x(\lambda) = \frac{1}{2\pi}\left|\frac{\sum_{l=0}^{q}\theta_l e^{-il\lambda}}{\sum_{k=0}^{p}\varphi_k e^{-ik\lambda}}\right|^2, \quad \lambda\in\Pi. \qquad (4.186)$$

将(4.185)代入(4.186)则得(4.184)。 ▎

利用(4.184)，ARMA 模型的谱估计可改为：

(a) 由样本$\{x_t\}_1^N$出发，先求$\hat{\rho}_0,\hat{\rho}_1,\cdots,\hat{\rho}_p,\cdots,\hat{\rho}_{q+p}$.

(b) 由延伸的 Yule-Walker 方程(4.174)解出$\{\hat{\varphi}_k\}_1^p$.

(c) 利用(4.175)计算出$\{\hat{\rho}_y(n)\}_0^q$.

(d) x_t 的 ARMA 谱密度估计为

$$\hat{f}_N(\lambda) = \frac{1}{2\pi}\frac{\hat{\rho}_y(0) + 2\sum_{l=1}^{q}\hat{\rho}_y(l)\cos l\lambda}{\left|\sum_{k=0}^{p}\hat{\varphi}_k e^{-ik\lambda}\right|^2}. \qquad (4.187)$$

因此，这一方法可以免去求(4.176)的非线性方程组。

§6 谱估计参数方法的 Bloomfield 指数模型

Bloomfield 在[56]中介绍了一种指数模型的谱估计方法。他的谱密度参数模型为

$$f(\lambda) = \frac{\sigma^2}{2\pi}e^{2\sum_{k=1}^{p}\gamma_k\cos k\lambda}, \quad \lambda\in\Pi, \qquad (4.188)$$

其中$\sigma^2,\{\gamma_k\}_1^p$为未知参数，这里$p$是假定已知的。

(1) 设x_t是平稳列，其谱密度具 Bloomfield 模型(4.188)，则x_t是正则的，即(参看(2.70))

$$\int_{\Pi}\log f_x(\lambda)d\lambda > -\infty. \qquad (4.189)$$

证明 由(4.188)，

$$\log f_x(\lambda) = \log\left(\frac{\sigma^2}{2\pi}\right) + 2\sum_{k=1}^{p} \gamma_k \cos k\lambda. \qquad (4.190)$$

由此

$$\int_{\Pi} \log f_x(\lambda)\,\mathrm{d}\lambda = 2\pi \log\left(\frac{\sigma^2}{2\pi}\right) + 2\sum_{k=1}^{p} \gamma_k \int_{\Pi} \cos k\lambda\ \mathrm{d}\lambda$$

$$= 2\pi \log\left(\frac{\sigma^2}{2\pi}\right) > -\infty. \qquad (4.191)$$

一步预测误差

$$\sigma_0^2 = 2\pi \exp\left\{\frac{1}{2\pi}\int_{\Pi} \log f_x(\lambda)\,\mathrm{d}\lambda\right\}$$

$$= 2\pi\{\sigma^2/2\pi\} = \sigma^2. \qquad (4.192)$$

(2) 在(4.188)中

$$\gamma_k = \frac{1}{2\pi}\int_{\Pi} \log f_x(\lambda) \cos k\lambda \,\mathrm{d}\lambda, \quad k = 1, 2, \cdots, p. \qquad (4.193)$$

证明 由(4.190)，

$$2\sum_{k=1}^{p} \gamma_k \cos k\lambda = \log f_x(\lambda) + \log\left(\frac{2\pi}{\sigma^2}\right). \qquad (4.194)$$

$$2\sum_{k=1}^{p} \gamma_k \cos k\lambda \cos l\lambda = \log\left(\frac{2\pi}{\sigma^2}\right)\cos l\lambda + \log f_x(\lambda) \cos l\lambda,$$

$$2\sum_{k=1}^{p} \gamma_k \int_{\Pi} \cos k\lambda \cos l\lambda\,\mathrm{d}\lambda$$

$$= \log\left(\frac{2\pi}{\sigma^2}\right)\int_{\Pi} \cos l\lambda\ \mathrm{d}\lambda + \int_{\Pi} (\log f_x(\lambda)) \cos l\lambda \,\mathrm{d}\lambda.$$

故得

$$\gamma_k = \frac{1}{2\pi}\int_{\Pi} (\log f_x(\lambda)) \cos k\lambda \,\mathrm{d}\lambda, \quad k = 1, 2, \cdots, p.$$

$$(4.195)$$

(3) 由样本出发的模型参数估计。

设 x_t 的样本为 $x_1, x_2, \cdots, x_N$。用 Bloomfield 模型(4.188) 来拟合 $f_x(\lambda)$。

首先，作为一种近似，先用$\{x_t\}_1^N$ 的周期图(见第五章)

$$I_N(\lambda)=\frac{1}{2\pi N}\left|\sum_{k=1}^{N} x_k e^{-ik\lambda}\right|^2, \quad \lambda\in\Pi \qquad (4.196)$$

来初步代替谱 $f_x(\lambda)$的位置，并对(4.195)取积分的分割和，$\Delta=\frac{2\pi}{N}$，则

$$\frac{1}{2\pi}\int_{\Pi}(\log f_x(\lambda))\cos l\lambda d\lambda=\gamma_l$$

可用

$$\frac{1}{\pi}\sum_{k}\log I_N(\lambda_k)\cos(l\lambda_k)\frac{2\pi}{N} \qquad (4.197)$$

来近似。当取$\lambda_k=\frac{2\pi}{N}k$ 时，求和范围为 $k=1,2,\cdots,\left[\frac{N-1}{2}\right]$。用

$$\hat{\gamma}_l=\frac{2}{N}\sum_{k=1}^{\left[\frac{N-1}{2}\right]}\left(\log I_N\left(\frac{2k\pi}{N}\right)\right)\cos\frac{2k\pi}{N}l, \quad l=1,2,\cdots,p \ (4.198)$$

来估计(4.188)中的参数$\{\gamma_k\}_1^p$。

以上对$\{\gamma_k\}_0^p$的估计在[37]中指出也是相容估计。

至于对 σ^2 的估计，[56]提出以

$$\hat{\sigma}^2=2\pi\exp\left\{\gamma+\frac{2}{N}\sum_{k=1}^{\left[\frac{N-1}{2}\right]}\log I_N\left(\frac{2k\pi}{N}\right)\right\} \qquad (4.199)$$

为估计，其中 $\gamma=0.57722\cdots$ 为 Euler 常数(见[57])。

近年来的文章进一步证实(4.199)估计的合理性，[58]中给出了如下关于(4.199)是强相容的定理：

定理4.12 设 x_t 是平稳正态序列，其谱密度 $f_x(\lambda)$满足

(1) $f_x(\lambda)>0$, $\lambda\in\Pi$.

(2) 对 $\alpha>\frac{1}{2}$，存在常数 $A>0$,使

$$\sup_{\lambda}|f_x(\lambda+\delta)-f_x(\lambda)|\leqslant A|\delta|^{\alpha},\tag{4.200}$$

则

$$\lim_{N\to\infty}\hat{\sigma}_N^2\triangleq\lim_{N\to\infty}\left(2\pi\exp\left\{\gamma+\frac{1}{2\pi}\int_{\Pi}\log I_N(\lambda)\mathrm{d}\lambda\right\}\right)\tag{4.201}$$

$$=\sigma_x^2,\quad \text{a.s}\tag{4.202}$$

其中 σ_x^2 为一步预测误差。

(4.201)括号的近似积分和正是(4.199)式。

Bloomfield 模型中的拟合阶数 p(见(4.188)) 假定是先验已知的。如果由样本出发对一个平稳列 x_t 进行 Bloomfield 指数模型的谱估计，则需要对阶数 p 进行可行的估计。

[37]中介绍了 Whittle 的近似对数似然函数是

$$L(x_t)=-\frac{N}{4\pi}\int_{\Pi}\left[\log(2\pi)^2f_x(\lambda)+\frac{I_N(\lambda)}{f_x(\lambda)}\right]\mathrm{d}\lambda\tag{4.203}$$

$$=-\frac{N}{4\pi}\left(\text{const}+\int_{\Pi}\left[\log f_x(\lambda)+\frac{I_N(\lambda)}{f_x(\lambda)}\right]\mathrm{d}\lambda\right).$$

取$f_x(\lambda)$为(4.188)的指数模型，则

$$\int_{\Pi}\log f_x(\lambda)\mathrm{d}\lambda=2\pi\log\left(\frac{\sigma^2}{2\pi}\right)+2\sum_{j=1}^{p}\gamma_j\int_{\Pi}\cos j\lambda\mathrm{d}\lambda$$

$$=2\pi\log\left(\frac{\sigma^2}{2\pi}\right).\tag{4.204}$$

故

$$L(x_t)=-\frac{N}{4\pi}\left(\text{const}+2\pi\log\left(\frac{\sigma^2}{2\pi}\right)\right.$$

$$\left.+\int_{\Pi}I_N(\lambda)\frac{2\pi}{\sigma^2}\,\mathrm{e}^{-2\sum\limits_{j=1}^{p}\gamma_j\cos j\lambda}\mathrm{d}\lambda\right)$$

$$= -N\,\text{const} - N\frac{1}{2\sigma^2}\int_{\Pi} I_N(\lambda)\exp\left\{-2\sum_{j=1}^{p}\gamma_j\cos j\lambda\right\}\mathrm{d}\lambda - N\log\sigma, \tag{4.205}$$

其中 const 是与诸参数无关的常数，第三项与 p 无关。我们记

$$L_N(p) = \frac{1}{2\sigma^2}\int_{\Pi} I_N(\lambda)\exp\left\{-2\sum_{j=1}^{p}\gamma_j\cos j\lambda\right\}\mathrm{d}\lambda. \tag{4.206}$$

而赤池的原本 AIC 公式是选一组极大似然参数估计 $\hat{\theta}^{(s)}$ 使

$$V(s) = -\log g(x\,|\,\hat{\theta}^{(s)}) + 2s \tag{4.207}$$

达极小，其中 $\log g(x\,|\,\theta)$ 为对数似然函数（参看(4.84)和(4.85)式），s 为模型参数个数。

考虑到(4.206)并且若选 $\hat{\sigma}^2$ 和 $\hat{\gamma}_j$ 为(4.199) 和(4.198)作为近似估计，则(4.207)可简化为求

$$\mathrm{AIC}(s) = \frac{1}{2\hat{\sigma}^2}\int_{\Pi} I_N(\lambda)\exp\left\{-2\sum_{j=1}^{s}\hat{\gamma}_j\cos j\lambda\right\}\mathrm{d}\lambda + \frac{2s}{N}, \tag{4.208}$$

$$0 < s < L(N)$$

的极小值 $p = s_0$ 作为阶的估计。

例4.1 取 x_t 为 ARMA(2,2)，模型方程为①

$$x_t - 0.2x_{t-1} - 0.1x_{t-2} = \varepsilon_t + 1.4\varepsilon_{t-1} + 0.5\varepsilon_{t-2}.$$

样本分别取 $N = 128, 256, 512$，得到的谱估计如图4.2所示。同时用极大熵得到的相应谱估计也作了比较（参看图4.3）。

为了便于计算，$\{\varepsilon_t\}$取为 i.i.d.$N(0,1)$，谱密度图只画正部。BL(p)表示 Bloomfield 拟合，阶估为 p（见(4.208)式）。

例4.2 x_t 为 MA(2)，模型方程为

---

① 这一部分读者可参看[83]。

$$x_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.7\varepsilon_{t-2}.$$

对 $N = 128, 256, 512$ 及和极大熵的比较示于图4.4及4.5.

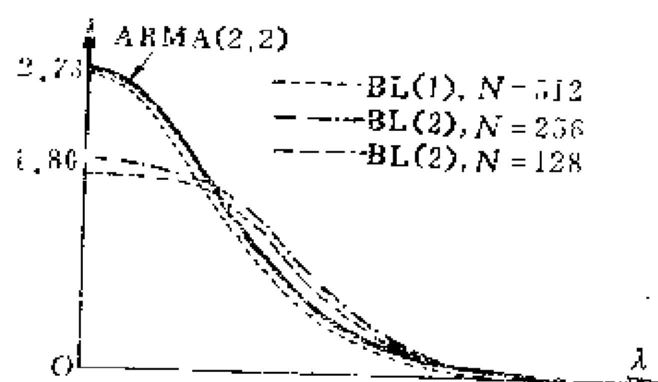


图 4.2

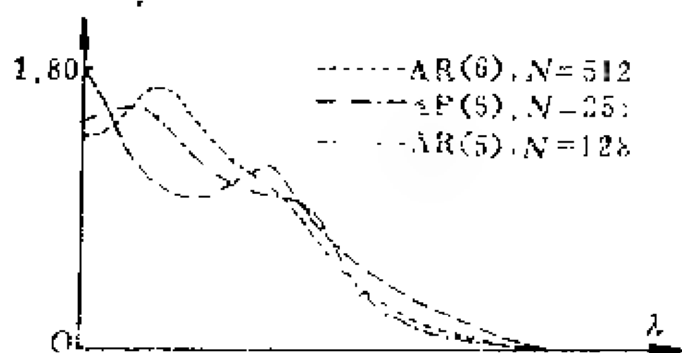


图 4.3

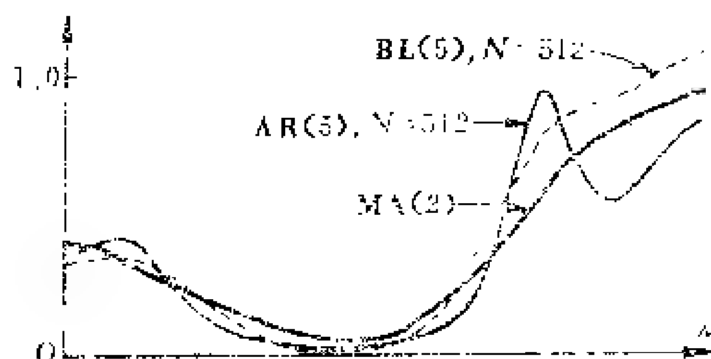


图 4.4

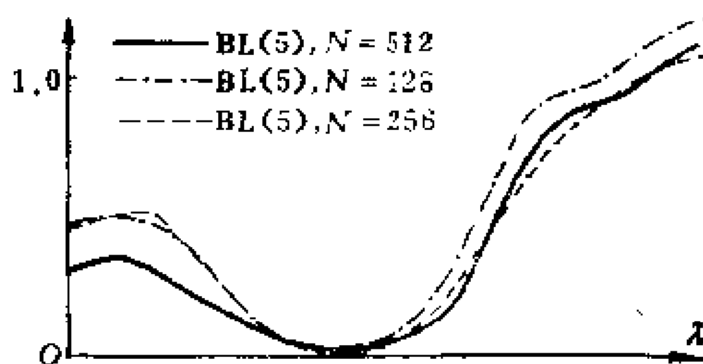


图 4.5

但 Bloomfield 模型也并非一切都好，如：

例4.3 x_t 取为 Beamish & Priestley(1981,[50]) 的 AR(4) 模型

$$x_t - 2.7607x_{t-1} + 3.8106x_{t-2} - 2.6535x_{t-3} + 0.9238x_{t-4} = \varepsilon_t.$$

当 $N = 256$ 时，Bloomfield 估计和极大熵估计的比较如图4.6，4.7 所示。

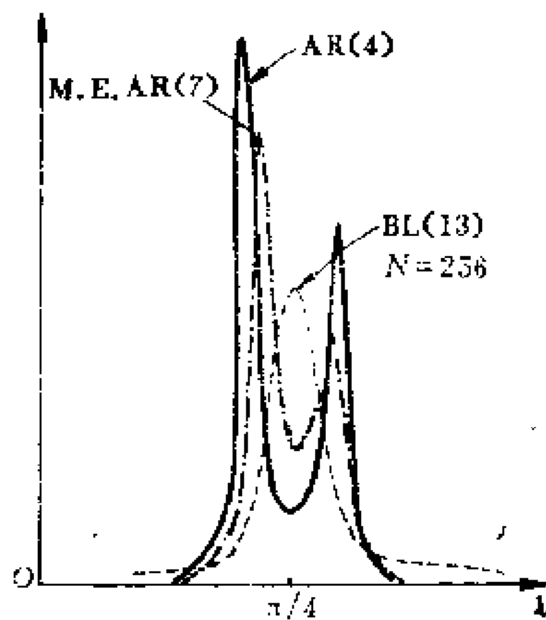


图 4.6

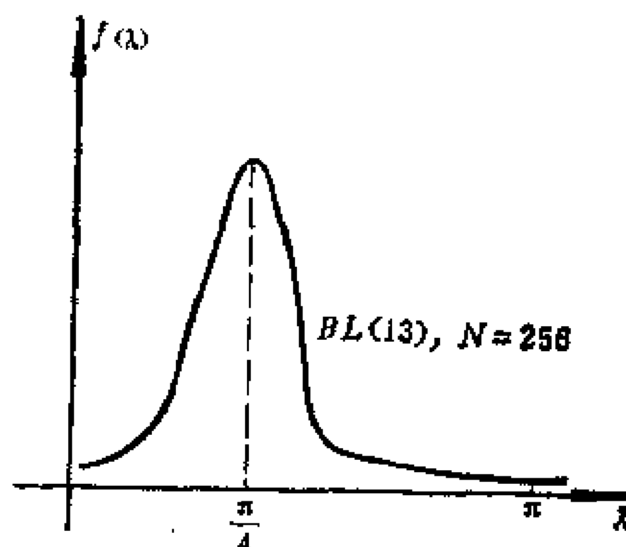


图 4.7

附录　关于判别代数多项式的根是否在单位圆外的方法

在高等代数(见[59],[60])中有 Sturm 序列的定义：

设 $f_0(x)$ 和 $f_1(x)$ 是两个 x 的多项式，其中 $f_1(x)$ 的次数低于 $f_0(x)$；用 $f_1(x)$ 去除 $f_0(x)$ 得商式 $q_0(x)$ 及余式 $R_2(x)$；$R_2(x)$ 的次数低于 $f_1(x)$。令

$$f_2(x) = -R_2(x)$$

为序列的第二个函数。如果 $f_2(x)\not\equiv 0$，则用 $f_2(x)$ 去除 $f_1(x)$ 得商式 $q_1(x)$ 及余式 $R_3(x)$，令

$$f_3(x) = -R_3(x)$$

为序列的第三个函数，…则可得 $\{f_0, f_1, \cdots, f_m\}$，而 $f_{m+1}\equiv 0$，于是称

$$\{f_0, f_1, f_2, \cdots, f_m\}$$

为以 f_0, f_1 为底的 Sturm 序列。

一般讲，Sturm 序列中三个相邻函数间的关系为

$$f_k(x) = q_k(x)f_{k+1}(x) - f_{k+2}(x),\quad k = 0, 1, 2, \cdots, m-1.$$

Routh 定理 设 $f(z)=z^n+a_1z^{n-1}+\cdots+a_n$ 为实系数多项式，$f(z)=0$ 在虚轴及右半平面上没有根的充要条件是：以

$$f_0(t)=t^n-a_2t^{n-2}+a_4t^{n-4}-\cdots \text{（有限项）}$$

及

$$f_1(t)=a_1t^{n-1}-a_3t^{n-3}+a_5t^{n-5}-\cdots \text{（有限项）}$$

为基的 Sturm 序列 $\{f_0, f_1, \cdots, f_s\}$ 内，每个多项式比它的前一个多项式低一次，$s=n$，而且首项系数都是正的。

Routh 定理解决了虚轴和右半平面上 n 阶实系数多项式的判根问题。我们可以作线性变换，把单位圆外变成 ω 复平面的右半平面。

设有实系数多项式

$$z^n+a_1z^{n-1}+\cdots+a_n=0.$$

线性变换

$$\omega=\frac{z+1}{z-1} \text{ 或 } z=\frac{\omega+1}{\omega-1}$$

把 z 平面的单位圆 $z\cdot\bar{z}=1$ 变成

$$\frac{(\omega+1)(\bar{\omega}+1)}{(\omega-1)(\bar{\omega}-1)}=\frac{1+\omega\bar{\omega}+(\omega+\bar{\omega})}{1+\omega\bar{\omega}-(\omega+\bar{\omega})}=1.$$

因此

$$2\mathrm{Re}(\omega)=\omega+\bar{\omega}=0,$$

即线性变换将圆周变成了虚轴，并把单位圆外变成了右半平面。

具体将上述 ω 变换代入多项式得

$$\left(\frac{\omega+1}{\omega-1}\right)^n+a_1\left(\frac{\omega+1}{\omega-1}\right)^{n-1}+\cdots+a_n=0$$

或

$$(\omega+1)^n+a_1(\omega+1)^{n-1}(\omega-1)+\cdots+a_n(\omega-1)^n=0.$$

若将上式表成

$$b(\omega)\triangleq b_0\omega^n+b_1\omega^{n-1}+\cdots+b_n=0,$$

则

$$B = TA,$$
$$B = (b_0, b_1, \cdots, b_n)',$$
$$A = (1,\ a_1, \cdots, a_n)',$$

T 是$(n+1)\times(n+1)$的方阵，

$$T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ c_n^1 & * & & & \\ c_n^2 & & t_{ij} & & \\ \vdots & & & & \\ c_n^n & & & & \end{bmatrix},$$

其中

$$t_{i,j} = t_{i,j-1} - t_{i-1,j-1} - t_{i-1,j}$$

（如 $* = c_n^1 - 1 - 1$）。

对 $b_0\omega^n + b_1\omega^{n-1} + \cdots + b_n = 0$ 用 Routh 判别法即可判明是否在虚轴和右半平面有根，即判 $f(z) = z^n + a_1z^{n-1} + \cdots + a_n$是否在单位圆上和外有根的问题，并可用来判别圆内外根的个数。

但是，有了 $b_0z^n + \cdots + b_n = 0$ 多项式如何用 Routh 定理求对应的 Sturm 序列仍不易。以下介绍 Routh 紧凑表格法（见[60]）。

首先把上述多项式系数写为首项为 1 的列$1, \hat{b}_1, \cdots, \hat{b}_n$（相应多项式记为 $\hat{b}(z)$），则以

$$f_0(t) = t^n - \hat{b}_2t^{n-2} + \hat{b}_4t^{n-4} - \cdots \qquad \text{（有限项）},$$
$$f_1(t) = \hat{b}_1t^{n-1} - \hat{b}_3t^{n-3} + \hat{b}_5t^{n-5} - \cdots \quad \text{（有限项）}$$

为基的 Sturm 序列$\{f_0, f_1, f_2, \cdots, f_n\}$求法如下（为明了起见，以下以$n=6$为例，列表）：

$$\begin{array}{l|l} f_0 & 1,\ \hat{b}_2,\ \hat{b}_4,\ \hat{b}_6 \\ f_1 & \hat{b}_1,\ \hat{b}_3,\ \hat{b}_5 \\ f_2 & D_0, D_1, D_2 \\ f_3 & E_0, E_1 \\ f_4 & F_0, F_1 \\ f_5 & G_0 \\ f_6 & F_1 \end{array}\left.\begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \end{array}\right\}$$

各数的计算式为 = (上二行右数) $-\dfrac{(\text{上二行最左数})\times(\text{上行右数})}{\text{上行最左数}}$

于是：

(1) 若 D_0, E_0, F_0, G_0, F_1 有零系数，则不满足首项系数皆为正的条件(Routh 定理)，因而 $\delta(z)$ 多项式在右半平面(包括虚轴)必有根。

(2) 若 $\delta(z)$ 系数有零或负数，必在右半平面(包括虚轴)上有根。

因为一对实部是负的共轭复根对应于

$$(z-(-a+i\beta))(z-(-a-i\beta))=z^2+2az+(a^2+\beta^2)$$

$$(a>0,\ \beta>0),$$

而 $a, a^2+\beta^2$ 皆为正。又：任何多个正系数多项式的乘积一定是正系数多项式。因此，如果 $\delta(z)$ 系数有零或负数，则必在右半平面(包括虚轴)有根。

例 1 $\delta(z)=z^4+5z^3+5z^2-5z-6$。由于有负系数 -5，故右(虚)半平面必有根。事实上可用 Routh 表格

$$\begin{array}{c|lll} f_0 & 1, & 5, & -6 \\ f_1 & 5, & -5 & \\ f_2 & 6, & -6 & \\ f_3 & 0 & & \end{array}$$

首项系数出现 0，故有根。事实上

$$\delta(z)=(z-1)(z+1)(z+2)(z+3)。$$

例 2 $\delta(z)=z^3+6z^2+11z+6$。

用 Routh 表格

$$\begin{array}{c|ll} f_0 & 1, & 11 \\ f_1 & 6, & 6 \\ f_2 & 10 & \\ f_3 & 6 & \end{array}$$

首项系数皆为正，故在右(虚)半面上无根。事实上

$$\hat{b}(z) = (z+1)(z+2)(z+3).$$

例3 试判 $\Phi(z) = 1 + 3z - 4z^2 + z^3$ 在单位圆内(上)是否有根?

首先记①: $\Phi^*(z) = z^3 + 3z^2 - 4z + 1$。作线性变换

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & -1 & -3 \\ 3 & -1 & -1 & 3 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 7 \\ -7 \end{bmatrix}.$$

故
$$b(z) = \hat{b}(z) = z^3 + 7z^2 + 7z - 7.$$

由于出现了负系数，故 $b(z)$ 在右(虚)半平面有根，从而 $\Phi^*(z)$ 在单位圆外(上)有根，即 $\Phi(z)$ 在单位圆内(上)有根。

习 题

1. 设 ξ 的概率密度为

$$p_\xi(x) = \begin{cases} \dfrac{\log 2}{x \log^2 x}, & x \geqslant 2; \\ 0, & \text{其它}. \end{cases}$$

但

$$H(\xi) = -\int_{R_1} p_\xi(x) \log p_\xi(x) dx$$

并不存在。

2. 在定理4.2中证明了：在给定二阶特性的条件下（一维场合是给定表征随机过程平均功率的方差 σ^2），使(4.4)的熵达最大的分布是正态分布。但如果条件改为只考虑在有限区间范围内取值的随机变量

① $\Phi^*(z) = z^3\left(1 + 3\frac{1}{z} - 4\frac{1}{z^2} + \frac{1}{z^3}\right) = z^3\Phi\left(\frac{1}{z}\right)$，因此 $\Phi^*(z)$ 在单位圆外（上）有根等价于 Φ 在单位圆内（上）有根。

$$p(x)=\begin{cases}p(x)>0, & a\leqslant x\leqslant b;\\ 0, & \text{其它}.\end{cases}$$

试证使 $H(\xi)$ 达极大的是均匀 $U[a,b]$ 分布。

3. 设 $\boldsymbol{\xi}=(\xi_1,\xi_2,\cdots,\xi_n)$，变换为 $\boldsymbol{\eta}=(\eta_1,\eta_2,\cdots,\eta_n)$，其中

$$\eta_i=f_i(\xi_1,\xi_2,\cdots,\xi_n),\quad i=1,2,\cdots,n.$$

则

$$H(\boldsymbol{\eta})=H(\boldsymbol{\xi})-\int_{R_n}p(\boldsymbol{x})\log\left|J\left(\frac{y_1,\cdots,y_n}{x_1,\cdots,x_n}\right)\right|\mathrm{d}\boldsymbol{x},$$

其中

$$J\left(\frac{y_1,\cdots,y_n}{x_1,\cdots,x_n}\right)=\begin{vmatrix}\frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \vdots & & \vdots\\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n}\end{vmatrix}.$$

但若 $\{f_i\}$ 为正交变换，则 $\boldsymbol{\xi}\to\boldsymbol{\eta}$ 的变换熵不变。

4. 唯一性问题：设给了 p 个实数 $\varphi_1,\varphi_2,\cdots,\varphi_p$，考虑满足线性方程组（Yule-Walker 方程形式）

$$\begin{bmatrix}R(0) & R(1) & \cdots & R(p-1)\\ \vdots & R(0) & \cdots & R(p-2)\\ \vdots & & \ddots & \vdots\\ R(p-1) & \cdots & \cdots & R(0)\end{bmatrix}\begin{bmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_p\end{bmatrix}=-\begin{bmatrix}R(1)\\ R(2)\\ \vdots\\ R(p)\end{bmatrix}$$

的 $\{R(0),R(1),\cdots,R(p)\}$。问它们是否唯一被确定？

试证：若由 $\{\varphi_1,\varphi_2,\cdots,\varphi_p\}$ 确定的

$$\Phi(z)=\sum_{l=0}^{p}\varphi_l z^l\neq 0,\quad |z|\leqslant 1,\ \varphi_0=1,$$

并且矩阵 $\boldsymbol{R}_p>0$（正定阵），则满足上述方程组的 $\{R(k)\}_0^p$ 在允许差一常数倍的意义下是唯一的。

5. 极大似然估计与 Kullback-Leibler 信息量的内在关系：

设总体的密度$p(x)$为具有以 $\boldsymbol{\theta}$ 为参数的分布族 $g(x|\boldsymbol{\theta})$ 取 $\boldsymbol{\theta}=\boldsymbol{\theta}_0$ 的函数，即

$$p(x)=g(x|\boldsymbol{\theta}_0).$$

设有来自总体的 n 个独立样本 $x_1,x_2,\cdots,x_n$，$\hat{\boldsymbol{\theta}}$ 是使对数似然函数

$$\frac{1}{n}L(\boldsymbol{\theta})$$

达极大的估计量。试说明：当 n 充分大，$\hat{\boldsymbol{\theta}}$和使得 K-L 信息量

$$I(p;\ g(\cdot|\tilde{\boldsymbol{\theta}}))=E_{\boldsymbol{\theta}_0}\left[\log\frac{g(\cdot|\boldsymbol{\theta}_0)}{g(\cdot|\tilde{\boldsymbol{\theta}})}\right]$$

达极小的参数估计$\tilde{\boldsymbol{\theta}}_0$是渐近相同的。

6. 设 x_t 是以 $(\varphi_1,\varphi_2,\cdots,\varphi_p;\ c_0,\theta_1,\theta_2,\cdots,\theta_q)=\boldsymbol{\theta}$ 为参数的正态 ARMA 模型

$$\sum_{k=0}^{p}\varphi_k x_{t-k}=c_0\left(\varepsilon_t+\sum_{l=1}^{q}\theta_l\varepsilon_{t-l}\right),\quad \varepsilon_t\sim N(0,1),$$

对样本$(x_1,x_2,\cdots,x_N)=\boldsymbol{X}$，记

$$E(\boldsymbol{X}'\boldsymbol{X})=c_0^2\boldsymbol{M}_N^{-1},$$

则正态 ARMA 模型 $\boldsymbol{X}$ 的对数似然函数为

$$L(\boldsymbol{\theta})=-\frac{N}{2}\ln(2\pi)-\frac{N}{2}\ln(c_0^2)+\frac{1}{2}\ln(\det\boldsymbol{M}_N)-\frac{1}{2c_0^2}\boldsymbol{X}\boldsymbol{M}_N\boldsymbol{X}'.$$

进而可进一步证明：设 $\hat{\boldsymbol{\theta}}$ 为极大似然估计，则

$$L(\hat{\boldsymbol{\theta}})=-\frac{N}{2}\ln(2\pi c_0)+\frac{1}{2}\ln(\det\boldsymbol{M}_N)-\frac{N}{2},$$

其中 c_0 为 $\hat{\boldsymbol{\theta}}$ 的分量。

7. 设 x_t 是正态 AR(p)列，$x_1,x_2,\cdots,x_N$是一段观测样本，令

$$\hat{\gamma}_j=\frac{1}{N}\sum_{l=1}^{N-j}x_l x_{l+j},\quad j=0,1,2,\cdots,N-1,$$

则

$$(1)\quad \boldsymbol{\Gamma}_{p+1}=\begin{bmatrix}\hat{\gamma}_0 & \hat{\gamma}_1 & \cdots & \hat{\gamma}_p\\ \vdots & \hat{\gamma}_0 & & \vdots\\ \vdots & & \ddots & \vdots\\ \hat{\gamma}_p & \cdots\cdots & & \hat{\gamma}_0\end{bmatrix}>0,\qquad \text{a.s.}.$$

(2) 由$(\hat{\gamma}_0,\hat{\gamma}_1,\cdots,\hat{\gamma}_p)$组成的 Yule-Walker 方程的$\hat{\boldsymbol{\theta}}$解$_N=(\boldsymbol{\varphi}_1,\cdots,\boldsymbol{\varphi}_p,\bar{c}_0)$(见(4.19))，当$N$充分大时有

$$\hat{\boldsymbol{\theta}}_N\to\boldsymbol{\theta}=(\varphi_1,\cdots,\varphi_p,c_0),\qquad \text{a.s.}.$$

其中$\boldsymbol{\theta}$为x_t AR(p)模型的眞参数，表明$\hat{\boldsymbol{\theta}}_N$是$\boldsymbol{\theta}$的强相容估计。

8. 设x_t为ARMA(0,0)，即白噪声$N(0,\sigma^2)$，又测得样本为$x_1,x_2,\cdots,x_N$，由第六题的极大似然函数表达式求极值可得

$$\hat{\sigma}^2=\frac{1}{N}\sum_{s=1}^{N}x_s^2.$$

第五章 谱估计的非参数方法

§1 引 言

上一章中较详细地讨论了谱估计中的参数方法。但从历史的回顾看，应该说发展谱分析的非参数方法比参数方法更早一些，这可以追溯到上个世纪末 Schuster(见[61])引用周期图作观测数据的周期性分析：设 $x_1, x_2, \cdots, x_N$ 是对观测对象的观测样本序列，称

$$I_N(\lambda) = \frac{1}{2\pi N}\left|\sum_{k=1}^{N} x_k e^{-ik\lambda}\right|^2, \quad -\pi \leqslant \lambda \leqslant \pi \tag{5.1}$$

为周期图。Schuster 利用 $I_N(\lambda)$的变化发现了许多气象、天文上的周期性(见[62])。从数学分析中的 Fourier 分析看，(5.1)有它的合理性，因为其中的

$$\sum_{k=1}^{N} x_k e^{-ik\lambda} \tag{5.2}$$

正是把$\{x_k\}_1^N$当成普通函数值的近似 Fourier 级数，而(5.1)相当于$\{x_k\}_1^N$的平均能谱(功率谱)。半个多世纪以来，周期图在天文、气象、地质、医学、经济、通讯、电子技术等极其广泛的领域中得到了成功的应用；此外，在理论上它也占有很突出的地位，如，由周期图引出的一个理论问题是它是否能充当随机序列谱密度的好的估计量。数学上证明了它是渐近无偏估计但却不是相容估计。于是，许多统计学家和电气工程师提出了一系列的理论问题——至今许多著名的数学家、统计学家、概率论专家还在利用改进的周期图寻找各种谱密度和谱函数的估计方法。因此，在某种意义上讲，周期图在非参数谱分析方法的理论和应用中乃是

占有核心地位的统计量。

本章将介绍以周期图为中心内容的谱估计理论和方法，它们不同于前一章，这些理论和方法的前提不假定有限参数模型。具体说来，由观测样本$\{x_k\}_1^N$为基本出发点将构造不同的统计量分别对：x_t的谱具有谱密度(绝对连续型)或x_t的谱是阶跃型(离散谱)或混合型的谱函数进行估计并从数学上证明这些估计量的合理性。

§2 周期图的基本统计分析

这一节主要讨论由(5.1)所定义的周期图在x_t(研究对象)为不同类型下的统计性质。先讨论x_t是具有谱密度的平稳列的情况。

1° 谱函数绝对连续条件下$I_N(\lambda)$的性质

定理5.1 设x_t是平稳列①，其相关函数$R_x(k)$满足

$$\sum_{k=-\infty}^{\infty}|R_x(k)|<+\infty, \tag{5.3}$$

$f_x(\lambda)$为其谱密度，则

$$\lim_{N\to\infty} EI_N(\lambda)=f_x(\lambda), \quad -\pi\leqslant\lambda\leqslant\pi, \tag{5.4}$$

其中$I_N(\lambda)$由(5.1)所定义。

证明 先将(5.1)改写为($\nu=k-j$)

$$I_N(\lambda)=\frac{1}{2\pi N}\sum_{j=1}^{N}\sum_{k=1}^{N}x_kx_je^{-i(k-j)\lambda} \qquad (5.5)'$$

$$=\frac{1}{2\pi N}\sum_{j=1}^{N}\left(\sum_{\nu=1-j}^{N-j}x_jx_{\nu+j}e^{-i\nu\lambda}\right)$$

---

① 同前几章一样，如不声明，即$Ex_t\equiv0$，x_t是实的。

$$= \frac{1}{2\pi N}\left\{\sum_{\nu=1-N}^{-1}\sum_{j=1-\nu}^{N} + \sum_{\nu=0}^{N-1}\sum_{j=1}^{N-\nu}\right\} x_j x_{\nu+j} e^{-i\lambda\nu}. \tag{5.5}$$

将大括号中第一项令 $j+\nu=l$，则

$$I_N(\lambda) = \frac{1}{2\pi N}\left(\sum_{\nu=1-N}^{-1}\sum_{l=1}^{N+\nu} x_l x_{l-\nu} + \sum_{\nu=0}^{N-1}\sum_{l=1}^{N-\nu} x_l x_{l+\nu}\right) e^{-i\nu\lambda}$$

$$= \frac{1}{2\pi N}\left(\sum_{\nu=1-N}^{-1}\sum_{l=1}^{N-|\nu|} + \sum_{\nu=0}^{N-1}\sum_{l=1}^{N-|\nu|}\right) x_l x_{l+|\nu|} e^{-i\nu\lambda}$$

$$= \frac{1}{2\pi N}\sum_{\nu=1-N}^{N-1}\left\{\sum_{l=1}^{N-|\nu|} x_l x_{l+|\nu|}\right\} e^{-i\nu\lambda},$$

即

$$I_N(\lambda) = \frac{1}{2\pi}\sum_{\nu=1-N}^{N-1} \hat{\gamma}_\nu e^{-i\nu\lambda}, \tag{5.6}$$

其中$\{\hat{\gamma}_\nu\}$和(4.81)一样是样本协方差

$$\hat{\gamma}_k = \frac{1}{N}\sum_{l=1}^{N-|k|} x_l x_{l+k}, \quad k = 0, \pm 1, \cdots, \pm(N-1). \tag{5.7}$$

由(5.6)

$$EI_N(\lambda) = \frac{1}{2\pi}\sum_{k=1-N}^{N-1} E[\hat{\gamma}_k] e^{-ik\lambda}$$

$$= \frac{1}{2\pi}\sum_{k=1-N}^{N-1}\left(1 - \frac{|k|}{N}\right) R_x(k) e^{-ik\lambda} \tag{5.8}$$

$$= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} R_N(k) e^{-ik\lambda}, \tag{5.9}$$

其中

$$R_N(k) = \begin{cases} \left(1 - \frac{|k|}{N}\right) R_x(k), & |k| = 0, 1, \cdots, N-1; \\ 0, & \text{其它。} \end{cases} \tag{5.10}$$

显见：

(1) $R_N(k) \to R_x(k), \quad N \to \infty.$

(2) $|R_N(k)|\leqslant|R_x(k)|$, $\sum_k |R_x(k)|<+\infty$.

故

$$\lim_{N\to\infty} EI_N(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}(\lim_{N\to\infty}R_N(k))e^{-ik\lambda}$$

$$=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}R_x(k)e^{-ik\lambda} \tag{5.11}$$

$$=f_x(\lambda).$$

其中最后一步是由 Wiener-Хинчин 定理得来。 ▎

如果我们对 $f_x(\lambda)$ 加强条件，如满足 Lipschitz 条件或具可微性等，则还可以估得(5.4)收敛的速度。

定理5.2 设 x_t 为平稳列，其谱密度 $f_x(\lambda)$ 在 $[-\pi,\pi]$ 上具有连续一级导数①，则

$$E(I_N(\lambda))=f_x(\lambda)+O\left(\frac{\log N}{N}\right),\quad -\pi\leqslant\lambda\leqslant\pi. \tag{5.12}$$

证明

$$\left|\sum_{k=1}^{N}e^{ik(\mu-\lambda)}\right|^2=\left|e^{i(\mu-\lambda)}\sum_{k=0}^{N-1}e^{ik(\mu-\lambda)}\right|^2$$

$$=\frac{1-e^{iN(\mu-\lambda)}}{1-e^{i(\mu-\lambda)}}\cdot\frac{1-e^{-iN(\mu-\lambda)}}{1-e^{-i(\mu-\lambda)}}$$

$$=\frac{2-(e^{iN(\mu-\lambda)}+e^{-iN(\mu-\lambda)})}{2-(e^{i(\mu-\lambda)}+e^{-i(\mu-\lambda)})}$$

$$=\frac{1-\cos N(\mu-\lambda)}{1-\cos(\mu-\lambda)}=\left(\frac{\sin\frac{N(\mu-\lambda)}{2}}{\sin\frac{\mu-\lambda}{2}}\right)^2. \tag{5.13}$$

---

① 在边点上理解为 $f_x(\lambda)$ 以 2π 为周期作延拓，仍具此性质，以后均作此理解。

则由(5.5)′ $EI_N(\lambda)=\frac{1}{2\pi N}\sum_{k,j=1}^{N}\left(\int_{\Pi}e^{i(k-j)\mu}f_x(\mu)d\mu\right)e^{-i(k-j)\lambda}$

$$=\frac{1}{2\pi N}\int_{\Pi}\left|\sum_{k=1}^{N}e^{ik(\mu-\lambda)}\right|^2 f_x(\mu)d\mu$$

$$=\frac{1}{2\pi N}\int_{\Pi}\left(\frac{\sin\frac{N(\mu-\lambda)}{2}}{\sin\frac{\mu-\lambda}{2}}\right)^2 f_x(\mu)d\mu. \quad (5.14)$$

又不难证明恒有

$$\frac{1}{2\pi N}\int_{\Pi}\left(\frac{\sin\frac{N(\lambda-\mu)}{2}}{\sin\frac{\lambda-\mu}{2}}\right)^2 d\mu=1. \quad (5.15)$$

利用(5.15)被积函数的周期性(μ 是变量)以及规定 $f_x(\mu)$ 作周期延拓，则可导出

$$EI_N(\lambda)-f_x(\lambda)=\frac{1}{2\pi N}\int_{-\pi+\lambda}^{\pi+\lambda}\left(\frac{\sin\frac{N(\lambda-\mu)}{2}}{\sin\frac{\lambda-\mu}{2}}\right)^2$$

$$\times(f_x(\mu)-f_x(\lambda))d\mu. \quad (5.16)$$

以下将证明：对 $-\pi\leqslant\lambda\leqslant\pi$，必有

$$|EI_N(\lambda)-f_x(\lambda)|\leqslant K\frac{\log N}{N}, \quad (5.17)$$

其中 K 是不依赖于 N 和 λ 的常数。以下记

$$b_N(\lambda)=EI_N(\lambda)-f_x(\lambda),$$

并取 $0<A<\frac{\pi}{2}$，由定理的条件

$$f_x(\lambda+u)-f_x(\lambda)=f'_x(\lambda+\theta u)u,\quad -\pi\leqslant\lambda\leqslant\pi, \quad (5.18)$$

其中 $|\theta|<1$，$|u|<A<\frac{\pi}{2}$。于是对稍大的 N 即可使以 λ 为中心以 A/N 为半径的邻域及以 A 为半径的邻域满足

$$\delta_{A/N}(\lambda) \subset \delta_A(\lambda)$$

（见图5.1），其中

$$\Delta_1 = \delta_{A/N}(\lambda) \subset \Delta_2 = \delta_A(\lambda),$$

$$\Delta_2^* = \Delta_2 \backslash \Delta_1.$$

利用(5.18)可将积分(5.16)表为三个积分：



$\Delta_1 = \{\mu:\ |\mu - \lambda| < A/N\}$
$\Delta_2 = \{\mu:\ |\mu - \lambda| < A\}$
$\Delta_3 = \{\mu:\ A < |\mu - \lambda|\}$

图 5.1

$$b_N(\lambda) = \frac{1}{2\pi N}\left\{\int_{|\mu-\lambda|<\frac{A}{N}} + \int_{\frac{A}{N}<|\mu-\lambda|<A} + \int_{A\leqslant|\mu-\lambda|}\right\}$$

$$\times\left(\frac{\sin\frac{N(\mu-\lambda)}{2}}{\sin\frac{\mu-\lambda}{2}}\right)^2 (f_x(\mu) - f_x(\lambda))\mathrm{d}\mu \qquad (5.19)$$

$$= \frac{1}{2\pi N}\left\{\int_{|u|<\frac{A}{N}} + \int_{\frac{A}{N}<|u|<A}\right\}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2 f'_x(\lambda+\theta u)u\mathrm{d}u$$

$$+ \frac{1}{2\pi N}\int_{A\leqslant|u|<\pi}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2 (f_x(\lambda+u) - f_x(\lambda))\mathrm{d}u$$

$$\leqslant \frac{|f'_x|_{\max}}{2\pi N}\left\{\int_{|u|<\frac{A}{N}} + \int_{\frac{A}{N}<|u|<A}\right\}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2 |u|\mathrm{d}u$$

$$+\frac{|f_x|_{max}}{\pi N}\int_{A\leqslant|u|\leqslant\pi}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2 du \tag{5.20}$$

记上式为 $I_1+I_2+I_3$。 (5.21)

以下先估 I_3。

由于
$$0<A\leqslant|u|\leqslant\pi,$$

故
$$\sin^2\frac{A}{2}\leqslant\sin^2\frac{u}{2}\leqslant1.$$

所以

$$I_3=\frac{|f_x|_{max}}{\pi N}\int_{A\leqslant|u|\leqslant\pi}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2 du$$

$$\leqslant\frac{|f_x|_{max}}{\pi N\left(\sin\frac{A}{2}\right)^2}\int_{A\leqslant|u|\leqslant\pi}du\leqslant\frac{\beta_3}{N}, \tag{5.22}$$

其中 β_3 是不依赖于 λ,N 的常数。

再估 I_2:

$$I_2=\frac{|f'_x|_{max}}{2\pi N}\int_{\frac{A}{N}\leqslant|u|<A}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2|u|du$$

$$\leqslant\frac{|f'_x|_{max}}{2\pi N}\int_{\frac{A}{N}\leqslant|u|<A}\frac{1}{\left(\frac{\sin u/2}{u/2}\right)^2}\frac{4du}{|u|}. \tag{5.23}$$

由于
$$\frac{|u|}{2}<\frac{A}{2}<\frac{\pi}{4}, \tag{5.24}$$

故 $\left(\frac{\sin\frac{u}{2}}{\frac{u}{2}}\right)^2$ 的最小值在 $\frac{A}{2}$ 点达到(见图5.2)。

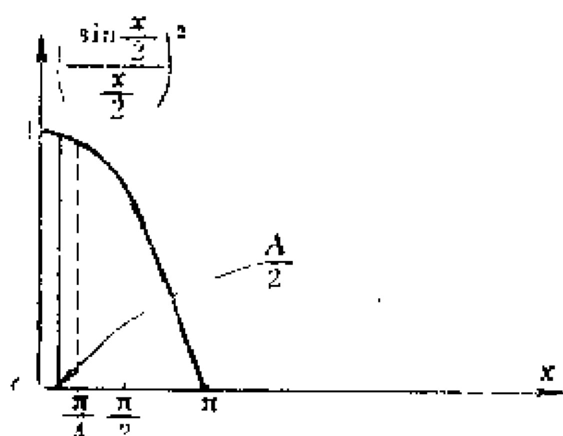


图 5.2

因此(5.23)右端可估为

$$I_2 \leqslant \frac{2|f'_x|_{\max}}{\pi N}\left(\frac{\frac{A}{2}}{\sin\frac{A}{2}}\right)^2 \int_{\frac{A}{N}<|u|<A}\frac{du}{|u|} = \beta_2\frac{\log N}{N}. \qquad (5.25)$$

同样 β_2 是与 N, λ 均无关的常数。

至于 I_1，可用以下方法估计：

$$I_1 = \frac{|f'_x|_{\max}}{2\pi N}\left\{\int_{|u|<\frac{A}{N}}\left(\frac{\sin\frac{Nu}{2}}{\sin\frac{u}{2}}\right)^2|u|du\right\}$$

$$= \frac{|f'_x|_{\max}}{2\pi N}\int_{|u|<\frac{A}{N}}\left(\frac{\sin\frac{Nu}{2}}{\frac{Nu}{2}}\right)^2\left(\frac{\sin\frac{u}{2}}{\frac{u}{2}}\right)^{-2}|u|N^2du$$

$$\leqslant \frac{N|f'_x|_{\max}}{2\pi}\int_{|u|<\frac{A}{N}}\left(\frac{\sin\frac{A}{2N}}{\frac{A}{2N}}\right)^{-2}|u|du$$

$$\leqslant \frac{A|f'_x|_{\max}}{2\pi}\left(\frac{\sin\frac{A}{2N}}{\frac{A}{2N}}\right)^{2}\int_{|u|<\frac{A}{N}}du. \qquad (5.26)$$

由于N充分大之后

$$\left(\frac{\sin\frac{A}{2N}}{\frac{A}{2N}}\right) > \frac{1}{2},$$

故

$$I_1 \leqslant \beta_1 / N, \tag{5.27}$$

其中 β_1 亦是不依赖于λ和N的常数。综合(5.27),(5.25)和(5.22)知(5.12)成立，并且 $O\left(\frac{\log N}{N}\right)$对$\lambda$还是一致的。 ▌

由定理5.1和5.2知：当 x_t 有谱密度并且适当满足一些条件时，周期图 $I_N(\lambda)$是 $f_x(\lambda)$谱密度的渐近无偏估计量。正是由于$I_N(\lambda)$具有这一好的性质才使得在工程技术界(包括气象、经济等部门中)广泛地应用它来作各种观测数据的周期性分析。但是，以下将指出：从统计分析上看，$I_N(\lambda)$并不是 $f_x(\lambda)$的好的估计量。例如它的方差 $D(I_N(\lambda)) \nrightarrow 0 (N\to\infty)$，因而一般说来 $I_N(\lambda)$在 L^2 意义下并不收敛到 $f_x(\lambda)$。在工程技术界中广泛流传的

$$\lim_{N\to\infty} I_N(\lambda) = f_x(\lambda) \tag{5.28}$$

一般讲是不成立的。

为此，我们在稍强一点的条件下来讨论这一问题。

定理5.3 设 x_t 是正态的平稳列，其谱密度 $f_x(\lambda)$在 Π 上具有连续一级导函数，则对 $\lambda,\mu\in[-\pi,\pi]$,

$$\operatorname{Cov}\{I_N(\lambda), I_N(\mu)\}$$

$$= \frac{f_x(\lambda)f_x(\mu)}{N^2}\left\{\frac{\sin^2\frac{N}{2}(\lambda+\mu)}{\sin^2\frac{1}{2}(\lambda+\mu)} + \frac{\sin^2\frac{N}{2}(\lambda-\mu)}{\sin^2\frac{1}{2}(\lambda-\mu)}\right\} + O\left(\frac{\log N}{N}\right). \tag{5.29}$$

证明 因证明较长，分为以下几步。

(1) 在正态条件下可得

$4\pi^2 N^2 \operatorname{Cov}(I_N(\lambda), I_N(\mu))$

$$= \int_{\Pi}\int_{\Pi}[D_N(\lambda+\alpha)D_N(\mu-\alpha)D_N(-\lambda+\beta)D_N(-\mu-\beta)$$
$$+ D_N(\lambda+\alpha)D_N(-\mu-\alpha)D_N(-\lambda+\beta)D_N(\mu-\beta)]f_x(\alpha)$$
$$\times f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta, \tag{5.30}$$

其中

$$D_N(\lambda) = \sum_{k=1}^{N} e^{ik\lambda} = e^{i\lambda}\frac{1-e^{iN\lambda}}{1-e^{i\lambda}}. \tag{5.31}$$

事实上，由

$$\mathrm{Cov}\left(\sum_i a_i\xi_i, \sum_j \beta_j\eta_j\right) = \sum_i\sum_j a_i\beta_j\,\mathrm{Cov}(\xi_i,\eta_j)$$

知

$$4\pi^2N^2\mathrm{Cov}(I_N(\lambda),I_N(\mu))$$
$$= \sum_{\nu_1,\nu_2,\nu_3,\nu_4=1}^{N} e^{i\nu_1\lambda}e^{-i\nu_2\lambda}e^{i\nu_3\mu}e^{-i\nu_4\mu}$$
$$\times\mathrm{Cov}(x_{\nu_1}x_{\nu_2},x_{\nu_3}x_{\nu_4}). \tag{5.32}$$

但在正态条件下

$$\mathrm{Cov}(x_{\nu_1}x_{\nu_2},x_{\nu_3}x_{\nu_4})$$
$$= E(x_{\nu_1}x_{\nu_2}x_{\nu_3}x_{\nu_4}) - E(x_{\nu_1}x_{\nu_2})E(x_{\nu_3}x_{\nu_4})$$
$$= R(\nu_1-\nu_2)R(\nu_3-\nu_4) + R(\nu_1-\nu_3)R(\nu_2-\nu_4)$$
$$+ R(\nu_1-\nu_4)R(\nu_2-\nu_3) - R(\nu_1-\nu_2)R(\nu_3-\nu_4)$$
$$= R(\nu_1-\nu_3)R(\nu_2-\nu_4) + R(\nu_1-\nu_4)R(\nu_2-\nu_3)$$
$$= \int_{\Pi}\int_{\Pi}(e^{i(\nu_1-\nu_3)\alpha}e^{i(\nu_2-\nu_4)\beta} + e^{i(\nu_1-\nu_4)\alpha}e^{i(\nu_2-\nu_3)\beta})$$
$$\times f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta. \tag{5.30*}$$

将上式代入(5.32)可得

$$4\pi^2N^2\mathrm{Cov}(I_N(\lambda),I_N(\mu))$$
$$= \int_{\Pi}\int_{\Pi}\sum_{j=1}^{N}\{e^{i\nu_1(\lambda-\alpha)}e^{i\nu_3(\mu-\alpha)}e^{-i\nu_2(\lambda-\beta)}e^{-i\nu_4(\mu+\beta)}$$
$$+ e^{i\nu_1(\lambda+\alpha)}e^{-i\nu_4(\mu+\alpha)}e^{-i\nu_2(\lambda-\beta)}e^{i\nu_3(\mu-\beta)}\}$$

$$\times f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta$$

$$=\int_\Pi\int_\Pi\{D_N(\lambda+\alpha)D_N(\mu-\alpha)D_N(-\lambda+\beta)D_N(-\mu-\beta)$$
$$+D_N(\lambda+\alpha)D_N(-\mu-\alpha)D_N(-\lambda+\beta)D_N(\mu-\beta)\}$$
$$\times f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta. \tag{5.33}$$

(2) 利用(5.13)和(5.31)的关系知

$$|D_N(\lambda)|^2=\left(\frac{\sin\frac{N}{2}\lambda}{\sin\frac{\lambda}{2}}\right)^2, \tag{5.34}$$

并且不难证明

$$\int_\Pi\int_\Pi(D_N(\lambda+\alpha)D_N(\mu-\alpha)D_N(-\lambda+\beta)D_N(-\mu-\beta)$$
$$+D_N(\lambda+\alpha)D_N(-\mu-\alpha)D_N(-\lambda+\beta)D_N(\mu-\beta))\mathrm{d}\alpha\mathrm{d}\beta$$
$$=(2\pi)^2(|D_N(\lambda+\mu)|^2+|D_N(\lambda-\mu)|^2).$$

因而可估 $(2\pi)^2(|D_N(\lambda+\mu)|^2+|D_N(\lambda-\mu)|^2)f_x(\lambda)f_x(\mu)$ 和(5.33)之差:

$$|4\pi^2N^2\mathrm{Cov}(I_N(\lambda),I_N(\mu))-4\pi^2(|D_N(\lambda+\mu)|^2+|D_N(\lambda-\mu)|^2)$$
$$\times f_x(\lambda)f_x(\mu)|$$
$$\leqslant\left|\int_\Pi\int_\Pi D_N(\lambda+\alpha)D_N(\mu-\alpha)D_N(-\lambda+\beta)D_N(-\mu-\beta)\right.$$
$$\left.\times(f_x(\alpha)f_x(\beta)-f_x(\lambda)f_x(\mu))\mathrm{d}\alpha\mathrm{d}\beta\right|$$
$$+\left|\int_\Pi\int_\Pi D_N(\lambda+\alpha)D_N(-\mu-\alpha)D_N(-\lambda+\beta)D_N(\mu-\beta)\right.$$
$$\left.\times(f_x(\alpha)f_x(\beta)-f_x(\lambda)f_x(\mu))\mathrm{d}\alpha\mathrm{d}\beta\right|$$
$$=I_1+I_2. \tag{5.33'}$$

由于 I_1 与 I_2 类似并且可得完全相同的阶,因而以下只需对 I_2 估计即可。由 Schwartz 不等式

$$I_2\leqslant\left[\int_\Pi\int_\Pi|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2|f_x(\alpha)f_x(\beta)\right.$$
$$-f_x(\lambda)f_x(\mu)|\mathrm{d}\alpha\mathrm{d}\beta$$

$$\times\int_{\Pi}\int_{\Pi}|D_N(-\mu-\alpha)|^2|D_N(-\lambda+\beta)|^2|f_x(\alpha)f_x(\beta)$$

$$-f_x(\lambda)f_x(\mu)|\mathrm{d}\alpha\mathrm{d}\beta\Big]^{1/2}$$

$$\triangleq[I_3I_4]^{1/2}. \tag{5.35}$$

由相似性质只需看 I_3 即可。

$$I_3=\int_{\Pi}\int_{\Pi}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2|f_x(\alpha)f_x(\beta)$$

$$-f_x(\lambda)f_x(\mu)|\mathrm{d}\alpha\mathrm{d}\beta$$

$$=\left(\iint_{E_1}+\iint_{E_2}\right)|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2\ |f_x(\alpha)f_x(\beta)$$

$$-f_x(\lambda)f_x(\mu)|\mathrm{d}\alpha\mathrm{d}\beta\triangleq I_{E_1}+I_{E_2}, \tag{5.36}$$

其中

$$\begin{aligned}E_1&=\{(\alpha,\beta):\ |\lambda+\alpha|\geqslant A\ \text{或}\ |\mu-\beta|\geqslant A\},\\ E_2&=\{(\alpha,\beta):\ |\lambda+\alpha|<A\ \text{且}\ |\mu-\beta|<A\},\end{aligned} \tag{5.37}$$

而 $0<A<\pi/2$。

$$I_{E_1}=\iint_{E_1}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2|f_x(\alpha)f_x(\beta)$$

$$-f_x(\lambda)f_x(\mu)|\mathrm{d}\alpha\mathrm{d}\beta$$

$$\leqslant K_1\iint_{E_1}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2\mathrm{d}\alpha\mathrm{d}\beta$$

$$=K_1\iint_{E_1}\left(\frac{\sin\frac{N}{2}(\lambda+\alpha)}{\sin\frac{1}{2}(\lambda+\alpha)}\right)^2\left(\frac{\sin\frac{N}{2}(\mu-\beta)}{\sin\frac{1}{2}(\mu-\beta)}\right)^2\mathrm{d}\alpha\mathrm{d}\beta. \tag{5.38}$$

由 E_1 的定义知，不妨设 $|\lambda+\alpha|\geqslant A$，则(5.38)的第一个被积函数有界记为 L_1，它是不依赖于 N 和 λ。

$$I_{E_1}\leqslant K_1L_1\iint_{E_1}\left(\frac{\sin\frac{N}{2}(\mu-\beta)}{\sin\frac{1}{2}(\mu-\beta)}\right)^2\mathrm{d}\alpha\mathrm{d}\beta$$

$$\leqslant K_2\int_E\left(\frac{\sin\dfrac{N}{2}(\mu-\beta)}{\sin\dfrac{1}{2}(\mu-\beta)}\right)^2 d\beta. \tag{5.39}$$

利用(5.15)，则

$$I_{E_1}\leqslant K_3N. \tag{5.40}$$

至于I_{E_2}，利用中值定理得

$$\begin{aligned}&f_x(\alpha)f_x(\beta)-f_x(\lambda)f_x(\mu)\\&\quad=f'_x(\bar\alpha)f_x(\bar\beta)(\alpha+\lambda)+f'_x(\bar\beta)f_x(\bar\alpha)(\beta-\mu),\end{aligned} \tag{5.41}$$

其中$(\bar\alpha,\bar\beta)$位于(α,β)到$(-\lambda,\mu)$的联线上，故

$$\begin{aligned}I_{E_2}=&\iint\limits_{E_2}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2\\&\times|f'_x(\bar\alpha)f_x(\bar\beta)(\alpha+\lambda)+f'_x(\bar\beta)f_x(\bar\alpha)(\beta-\mu)|d\alpha d\beta\end{aligned} \tag{5.42}$$

但由E_2区域性质(见(5.37))及(5.15)可估

$$\begin{aligned}I_{E_2}\leqslant&\max(|f'_x(u)||f_x(v)|)\\&\times\left[\iint\limits_{E_2}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2|\alpha+\lambda|d\alpha d\beta\right.\\&\left.+\iint\limits_{E_2}|D_N(\lambda+\alpha)|^2|D_N(\mu-\beta)|^2|\beta-\mu|d\alpha d\beta\right]\end{aligned} \tag{5.43}$$

$$\begin{aligned}\leqslant&C_1\int_{|\mu-\beta|<A}\left(\frac{\sin^2\dfrac{N}{2}(\mu-\beta)}{\sin^2\dfrac{1}{2}(\mu-\beta)}\right)d\beta\\&\times\int_{|\lambda+\alpha|<A}\left(\frac{\sin^2\dfrac{N}{2}(\lambda+\alpha)}{\sin^2\dfrac{1}{2}(\lambda+\alpha)}\right)|\alpha+\lambda|d\alpha\\&+C_1\int_{|\lambda+\alpha|<A}\left(\frac{\sin^2\dfrac{N}{2}(\lambda+\alpha)}{\sin^2\dfrac{1}{2}(\lambda+\alpha)}\right)d\alpha\end{aligned}$$

$$\times \int_{|\mu-\beta|<A} \left(\frac{\sin^2 \frac{N}{2}(\mu-\beta)}{\sin^2 \frac{1}{2}(\mu-\beta)} \right) |\beta-\mu| \mathrm{d}\beta$$

$$\leqslant C_2 N \int_{|\lambda+\alpha|<A} \left(\frac{\sin^2 \frac{N}{2}(\lambda+\alpha)}{\sin^2 \frac{1}{2}(\lambda+\alpha)} \right) |\alpha+\lambda| \mathrm{d}\alpha$$

$$+ C_2 N \int_{|\mu-\beta|<A} \left(\frac{\sin^2 \frac{N}{2}(\mu-\beta)}{\sin^2 \frac{N}{2}(\mu-\beta)} \right) |\mu-\beta| \mathrm{d}\beta, \qquad (5.44)$$

其中 C_1, C_2 不依赖于 λ 和 μ, N。比较(5.44)二个积分和定理 5.2 对 I_1, I_2 的估计知两者完全相似①，因而有相同的阶，即

$$I_{E_2} \leqslant C_3 N \log N. \qquad (5.45)$$

可见

$$I_3 = I_{E_1} + I_{E_2} \leqslant KN \log N. \qquad (5.46)$$

类似地对 I_4 也有 $K'N \log N$ 的阶的估计，从而

$$I_2 \leqslant (I_3 I_4)^{\frac{1}{2}} = O(N \log N). \qquad (5.47)$$

I_1 和 I_2 类似，有相同的阶②。回到(5.34)，(5.33′)知

$$4\pi^2 N^2 \mathrm{Cov}(I_N(\lambda), I_N(\mu))$$

$$= 4\pi^2 \left(\frac{\sin^2 \frac{N}{2}(\lambda+\mu)}{\sin^2 \frac{1}{2}(\lambda+\mu)} + \frac{\sin^2 \frac{N}{2}(\lambda-\mu)}{\sin^2 \frac{1}{2}(\lambda-\mu)} \right) f_x(\lambda) f_x(\mu)$$

$$+ O(N \log N), \qquad (5.48)$$

其中 $O(N \log N)$ 对 λ, μ 是一致的。再将 N^2 移到(5.48)的右边即证完定理的结论。 ▍

推论 在定理 5.3 的条件下，

---

① 二者差一个因子 N^2。

② 由定理5.2的证明可看出 (5.47) 对λ, μ是一致的。

(1) $$\lim_{N\to\infty} D(I_N(\lambda)) = \begin{cases} f_x^2(\lambda), & \lambda \neq 0, \pm\pi; \\ 2f_x^2(0) & \lambda = 0, \pm\pi. \end{cases} \tag{5.49}$$

(2) 当 $|\lambda \pm \mu| \neq 0, \pm\pi$ 时

$$\lim_{N\to\infty} \operatorname{Cov}\{I_N(\lambda), I_N(\mu)\} = 0. \tag{5.50}$$

证明 当 $\lambda = \mu$ 但不等于 $0, \pm\pi$ 时，由(5.29)

$$D(I_N(\lambda)) = \frac{f_x^2(\lambda)}{N^2}\{|D_N(2\lambda)|^2 + |D_N(0)|^2\} + O\left(\frac{\log N}{N}\right). \tag{5.51}$$

对给定 λ,

$$|D_N(2\lambda)|^2 = \left(\frac{\sin^2 N\lambda}{\sin^2\lambda}\right) \leqslant \frac{1}{\sin^2\lambda}, \tag{5.52}$$

$$|D_N(0)|^2 = N^2, \tag{5.53}$$

故 $D(I_N(\lambda)) \to f_x^2(\lambda)$(当 $N \to \infty$)，即为(5.49)的第一式。

又若 $\lambda = 0, \pm\pi$，则

$$D(I_N(\lambda)) = \frac{f_x^2(0)}{N^2}\{2N^2\} + O\left(\frac{\log N}{N}\right).$$

故 $D(I_N(\lambda)) \to 2f_x^2(0)$，即证得(5.49)式。

再则，若 $|\lambda \pm \mu| \neq 0, \pm\pi$，则

$$|D_N(\lambda + \mu)|^2 = \left(\frac{\sin^2 \frac{N}{2}(\lambda + \mu)}{\sin^2 \frac{1}{2}(\lambda + \mu)}\right)$$

及 $|D_N(\lambda - \mu)|^2$ 皆为有界量，故运用定理5.3

$$\lim_{N\to\infty} \operatorname{Cov}\{I_N(\lambda), I_N(\mu)\} = 0. \quad \blacksquare$$

由定理5.3及定理5.2知：当 x_t 的谱密度满足一定条件时，周期图 $I_N(\lambda)$ 虽是 $f_x(\lambda)$ 的渐近无偏估计量但却不是相容估计，即 $D(I_N(\lambda)) \nrightarrow 0$(当 $f_x(\lambda) \neq 0$)。又

$$E[I_N(\lambda) - f_x(\lambda)]^2 = E[(I_N(\lambda) - EI_N(\lambda)) + (EI_N(\lambda) - f_x(\lambda)]^2$$
$$= D(I_N(\lambda)) + b_N^2(\lambda) \tag{5.54}$$

而 $b_N(\lambda) \to 0$，但对 $f_x(\lambda) \neq 0$ 的那些点知

$$E[I_N(\lambda)-f_x(\lambda)]^2 \nrightarrow 0 \quad (N\to\infty), \tag{5.55}$$

表明 $I_N(\lambda)\overset{L_2}{\nrightarrow}f_x(\lambda)$。改进这一缺点的办法将在§3讨论。

2° 谱函数非绝对连续条件下 $I_N(\lambda)$ 的性质

为了更一般地来讨论周期图 $I_N(\lambda)$ 的性质，以下先讨论弱 P 平稳列。

定义5.1 设 ξ_t 是实随机序列，称它是四阶平稳的假若对任何整数 n_1,n_2,n_3 和 h 有

(1) $E[\xi_h]^4<+\infty$；

(2) $E[\xi_h\xi_{n_1+h}]=E[\xi_0\xi_{n_1}]$； (5.56)

(3) $E[\xi_h\xi_{n_1+h}\xi_{n_2+h}]=E[\xi_0\xi_{n_1}\xi_{n_2}]$； (5.57)

(4) $E[\xi_h\xi_{n_1+h}\xi_{n_2+h}\xi_{n_3+h}]=E[\xi_0\xi_{n_1}\xi_{n_2}\xi_{n_3}]$。 (5.58)

定义5.2 称实随机序列 ξ_t 满足弱 P 条件，假若

(1) ξ_t 是四阶平稳列；

(2) 对整数 n_1,n_2，

$$\sup_{n_1,n_2}\left\{\sum_{n=-\infty}^{\infty}|Q(n_1,n,n_2+n)|\right\}\leqslant M^{①}<+\infty, \tag{5.59}$$

其中

$$Q(n_1,n_2,n_3)=E(\xi_0\xi_{n_1}\xi_{n_2}\xi_{n_3})-E(\xi_0\xi_{n_1})E(\xi_{n_2}\xi_{n_3})$$
$$-E(\xi_0\xi_{n_2})E(\xi_{n_1}\xi_{n_3})-E(\xi_0\xi_{n_3})E(\xi_{n_1}\xi_{n_2}). \tag{5.60}$$

由上一段的(5.30)可以看出，如果 ξ_t 是正态平稳列，则 $Q(n_1,n_2,n_3)\equiv 0$，因而(5.59)必满足，所以弱 P 条件具有相当广泛的应用背景。事实上(见[63])：

(a) 当 ξ_t 为正态平稳列时必满足弱P条件(事实上为强P)。

---

① 数学上称 $\sum_{n_1,n_2,n_3=-\infty}^{+\infty}|Q(n_1,n_2,n_3)|<+\infty$ 为四阶强P条件。显见，强 P 必为弱 P。

(b) 当 ξ_t 是具有4阶矩的i.i.d.序列，则 ξ_t 满足弱 P(实为强 P)。

(c) 当 ξ_t 是具有4阶矩的实平稳列，如果它是 m 步相依的，即对整数

$$k_1<k_2<\cdots<k_s<(k_s)+m+1\leqslant l_1<l_2<\cdots<l_t \quad (5.61)$$

$(\xi_{k_1},\xi_{k_2},\cdots,\xi_{k_s})$ 与 $(\xi_{l_1},\xi_{l_2},\cdots,\xi_{l_t})$ 是相互独立的，则 ξ_t 满足弱 P(实为强 P)。

(d) 若 x_t 为ARMA序列，它的Wold序列 $\{\varepsilon_t\}$ 是满足(b)的，则 x_t 亦是强 P 序列。

此外若 ξ_t 是缺羡序列并满足适当条件(见[63])，则可满足强 P 条件因而更满足弱 P 条件。

定理5.4 设 ξ_t 为满足弱 P 条件的平稳列，谱密度存在，并 $f_\xi(\lambda)\in L^2(d\lambda)$。记 $B_\xi(\tau)$ 为相关函数，则对给定的 k，

$$Y_k(n)=\xi(n+k)\xi(n)-B_\xi(k) \quad (5.62)$$

为平稳列，且谱密度 $f_{Y_k}(\lambda)$ 满足如下的不等式

$$\sup_{\lambda,k} f_{Y_k}(\lambda)\leqslant 4\pi\int_B f_\xi^2(\lambda)d\lambda+M. \quad (5.63)$$

证明

$$EY_k(n)=E(\xi_{n+k}\xi_n)-B_\xi(k)=0. \quad (5.64)$$

$$\begin{aligned}
B_{Y_k}(n)&=E(Y_k(n+h)Y_k(h))\\
&=E[\xi_{h+k}\xi_h-B_\xi(k)][\xi_{n+h+k}\xi_{n+h}-B_\xi(k)]\\
&=E(\xi_h\xi_{n+h}\xi_{h+k}\xi_{n+h+k})-B_\xi^2(k)\\
&=E(\xi_0\xi_k\xi_n\xi_{n+k})-E(\xi_0\xi_k)E(\xi_n\xi_{n+k})\\
&\quad-E(\xi_0\xi_n)E(\xi_k\xi_{n+k})-E(\xi_0\xi_{n+k})E(\xi_k\xi_n)\\
&\quad+[E(\xi_0\xi_k)E(\xi_n\xi_{n+k})+E(\xi_0\xi_n)E(\xi_k\xi_{n+k})\\
&\quad+E(\xi_0\xi_{n+k})E(\xi_k\xi_n)]-B_\xi^2(k)\\
&=Q(k,n,n+k)+[B_\xi(k)B_\xi(k)+B_\xi(n)B_\xi(n)\\
&\quad+B_\xi(n+k)B_\xi(n-k)]-B_\xi^2(k)
\end{aligned}$$

$$= Q(k,n,n+k) + B_\xi^2(n) + B_\xi(n+k)B_\xi(n-k). \tag{5.65}$$

故

$$\sum_{n=-\infty}^{\infty} B_{Y_k}(n) = \sum_{n=-\infty}^{\infty} Q(k,n,n+k) + \sum_{n=-\infty}^{\infty} B_\xi^2(n) + \sum_{n=-\infty}^{\infty} B_\xi(n+k)B_\xi(n-k). \tag{5.66}$$

可见

$$\sum_{n=-\infty}^{\infty} |B_{Y_k}(n)| \leqslant \sum_{n=-\infty}^{\infty} |Q(k,n,n+k)| + \sum_{n=-\infty}^{\infty} (B_\xi(n))^2 + \sum_{n=-\infty}^{\infty} |B_\xi(n+k)||B_\xi(n-k)|. \tag{5.67}$$

利用弱P条件，则

$$\sum_{n=-\infty}^{\infty} |B_{Y_k}(n)| \leqslant M + 2\pi\int_\Pi f_\xi^2(\lambda)\mathrm{d}\lambda + \left[\sum_{l=-\infty}^{\infty}(B_\xi(l))^2 \sum_{l=-\infty}^{\infty}(B_\xi(l-2k))^2\right]^{\frac{1}{2}}$$

$$= 4\pi\int_\Pi f_\xi^2(\lambda)\mathrm{d}\lambda + M. \tag{5.68}$$

$$\sup_{k,\lambda} f_{Y_k}(\lambda) \leqslant \sup_k \sum_{n=-\infty}^{\infty} |B_{Y_k}(n)| \leqslant 4\pi\int_\Pi f_\xi^2(\lambda)\mathrm{d}\lambda + M. \tag{5.69}$$

定理5.5 设ξ_n为弱P平稳列，$f_\xi(\lambda)\in L^2(\mathrm{d}\lambda)$。则

(1) 对一切N，当$k\geqslant\frac{6}{8}$时，

$$E\left[\sup_\lambda \frac{1}{N^{2k}}\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|^2\right] = O\left(\frac{1}{N^{2k-\frac{3}{2}}}\right). \tag{5.70}$$

(2) $k > \frac{7}{8}$时

$$\sup_{\lambda} \left| \sum_{n=1}^{N} \xi_n e^{in\lambda} \right| = o(N^k), \quad \text{a.s.} . \tag{5.71}$$

证明 先证(5.70)。

$$\begin{aligned}
\left| \sum_{n=1}^{N} \xi_n e^{in\lambda} \right|^2 &= \sum_{l=1}^{N} \sum_{n=1}^{N} \xi_n \xi_l e^{i(n-l)\lambda} \\
&= \sum_{l=1}^{N} \sum_{\tau=1-l}^{N-l} \xi_l \xi_{l+\tau} e^{i\tau\lambda} \\
&= \sum_{\tau=1}^{N-1} \sum_{l=1}^{N-\tau} \xi_l \xi_{l+\tau} e^{i\tau\lambda} + \sum_{\tau=1-N}^{0} \sum_{l=1-\tau}^{N} \xi_l \xi_{l+\tau} e^{i\tau\lambda} \\
&= \sum_{j=0}^{N-1} \sum_{u=1}^{N-j} \xi_{u+j} \xi_u e^{-ij\lambda} + \sum_{\tau=1}^{N-1} \sum_{l=1}^{N-\tau} \xi_l \xi_{l+\tau} e^{i\tau\lambda} \\
&\leqslant \sum_{j=0}^{N-1} \left| \sum_{u=1}^{N-j} \xi_{u+j} \xi_u \right| + \sum_{\tau=1}^{N-1} \left| \sum_{l=1}^{N-\tau} \xi_{l+\tau} \xi_l \right| \\
&\leqslant 2 \sum_{\tau=0}^{N-1} \left| \sum_{l=1}^{N-\tau} \xi_{l+\tau} \xi_l \right|
\end{aligned}$$

与 λ 无关。

$$\begin{aligned}
&\sup_{\lambda} \left| \sum_{n=1}^{N} \xi_n e^{in\lambda} \right|^2 \\
&= 2 \sum_{\tau=0}^{N-1} \left| \sum_{l=1}^{N-\tau} [\xi_{l+\tau} \xi_l - B_\xi(\tau)] + [(N-\tau) B_\xi(\tau)] \right| \\
&\leqslant 2 \sum_{\tau=0}^{N-1} (N-\tau) |B_\xi(\tau)| + 2 \sum_{\tau=0}^{N-1} \left| \sum_{l=1}^{N-\tau} [\xi_{l+\tau} \xi_l - B_\xi(\tau)] \right|,
\end{aligned} \tag{5.72}$$

其中

$$2\sum_{\tau=0}^{N-1}(N-\tau)|B_\xi(\tau)|\leqslant 2\left(\sum_{\tau=0}^{N-1}(N-\tau)^2\right)^{\frac{1}{2}}\left(\sum_{\tau=0}^{N-1}|B_\xi(\tau)|^2\right)^{\frac{1}{2}}$$

$$\leqslant 2N^{\frac{3}{2}}\left(\sum_{\tau=0}^{\infty}|B_\xi(\tau)|^2\right)^{\frac{1}{2}}$$

$$\leqslant 2N^{\frac{3}{2}}\left(2\pi\int_\Pi f_\xi^2(\lambda)\mathrm{d}\lambda\right)^{\frac{1}{2}}$$

$$\leqslant C_1N^{\frac{3}{2}}. \tag{5.73}$$

又利用定理5.4得

$$E\left|\sum_{l=1}^{N-\tau}[\xi_{l+\tau}\xi_l-B_\xi(l)]\right|$$

$$=E\left|\sum_{l=1}^{N-\tau}Y_l(\tau)\right|\leqslant\sqrt{E\left|\sum_{l=1}^{N-\tau}Y_l(\tau)\right|^2}$$

$$=\sqrt{E\left|\int_\Pi\left(\sum_{l=1}^{N-\tau}e^{il\lambda}\mathrm{d}Z_{Y_\tau}(\lambda)\right|^2} \tag{5.74}$$

$$=\left(\int_\Pi\left|\sum_{l=1}^{N-\tau}e^{il\lambda}\right|^2 f_{Y_\tau}(\lambda)\mathrm{d}\lambda\right)^{\frac{1}{2}}$$

$$\leqslant\left(\sup_\lambda f_{Y_\tau}(\lambda)\right)^{\frac{1}{2}}\left(\int_\Pi\left|\sum_{l=1}^{N-\tau}e^{il\lambda}\right|^2\mathrm{d}\lambda\right)^{\frac{1}{2}}$$

$$=(\sup_\lambda\ f_{Y_\tau}(\lambda))^{\frac{1}{2}}(2\pi(N-\tau))^{\frac{1}{2}}$$

$$\leqslant C_2N^{\frac{1}{2}}. \tag{5.75}$$

于是

$$E\left(\sup_\lambda\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|^2\right)$$

$$\leqslant C_1N^{\frac{3}{2}}+2\sum_{\tau=0}^{N-1}(C_2N^{\frac{1}{2}})$$

$$\leqslant C_3 N^{\frac{3}{2}},$$

并且

$$E\left(\sup_{\lambda}\frac{1}{N^{2k}}\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|^2\right)\leqslant C_3 N^{\frac{3}{2}-2k}<+\infty. \qquad (5.76)$$

最后一个不等式是因 $k\geqslant\frac{3}{4}$。故定理的第一部分得证。

以下证第二部分。记

$$S(N)=N^{-k}\sup_{\lambda}\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|. \qquad (5.77)$$

则当 $k>\frac{7}{8}$ 时，存在一个正数 $\beta>1$ 满足

$$k-\frac{3}{4}>\frac{1}{2\beta}>\frac{1}{8}. \qquad (5.78)$$

(1) 当 $N=[M^\beta]$①，则由前面证的结果

$$\begin{aligned}E|S(N)|^2&=E|S([M^\beta])|^2\\&=E\left[N^{-k}\sup_{\lambda}\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|\right]^2\\&=E\left[N^{-2k}\sup_{\lambda}\left|\sum_{n=1}^{N}\xi_n e^{in\lambda}\right|^2\right]\\&\leqslant C\left([M^\beta]\right)^{\frac{3}{2}-2k}\ (\text{见}(5.76))\\&\leqslant C'(M)^{\beta\left(\frac{3}{2}-2k\right)}=C'M^{-\alpha}.\end{aligned} \qquad (5.79)$$

但由(5.78) $\alpha=\beta\left(2k-\frac{3}{2}\right)>1$，可见

$$\sum_{M=1}^{\infty}E|S([M^\beta])|^2<+\infty. \qquad (5.80)$$

① $[x]$表示 x 的整数部分。

因此

$$\sum_{M=1}^{\infty} P\{|S([M^\beta])| \geqslant \varepsilon\}$$

$$\leqslant \sum_{M=1}^{\infty} \frac{1}{\varepsilon^2} E|S([M^\beta])|^2 < +\infty. \tag{5.81}$$

可见

$$\lim_{M\to\infty} S([M^\beta]) = 0, \quad \text{a.s.}. \tag{5.82}$$

(2) 对一般的N序列，$N\to\infty$，可用第一章对定理1.16证明的手法，设 $N_M \leqslant N \leqslant N_{M+1}$，其中

$$N_M = [M^\beta]. \tag{5.83}$$

$$S(N) = N^{-k} \sup_\lambda \left|\left(\sum_{n=1}^{[M^\beta]} + \sum_{n=[M^\beta]+1}^{N}\right) \xi_n e^{in\lambda}\right|^{①}$$

$$\leqslant N^{-k}\left\{\sup_\lambda \left|\sum_{n=1}^{[M^\beta]} \xi_n e^{in\lambda}\right| + \sup_\lambda \left|\sum_{n=[M^\beta]+1}^{N} \xi_n e^{in\lambda}\right|\right\}$$

$$= N^{-k}[M^\beta]^k S([M^\beta]) + N^{-k} \sup_\lambda \left|\sum_{n=[M^\beta]+1}^{N} \xi_n e^{in\lambda}\right|.$$

$$|S(N) - N^{-k}[M^\beta]^k S([M^\beta])| \leqslant [M^\beta]^{-k} \sup_\lambda \left|\sum_{n=[M^\beta]+1}^{N} \xi_n e^{in\lambda}\right|. \tag{5.84}$$

从而考虑充分大的M

$$E\{\sup_{[M^\beta]\leqslant N\leqslant[(M+1)^\beta]} |S(N) - N^{-k}[M^\beta]^k S([M^\beta])|^2\}$$

$$\leqslant [M^\beta]^{-2k}\ E\left\{\sup_\lambda\left(\sum_{n=[M^\beta]+1}^{[(M+1)^\beta]} |\xi_n|\right)^2\right\}$$

$$\leqslant (M-1)^{-2k\beta} E\left\{\sum_{k,l=[M^\beta]+1}^{[(M+1)^\beta]} |\xi_k||\xi_l|\right\}$$

---

① 若 $N=[M^\beta]$则理解为无第二求和项。

$$\leqslant (M-1)^{-2k\beta} B_\xi(0)([(M+1)^\beta]-[M^\beta])^2. \tag{5.85}$$

然而$[(M+1)^\beta]-[M^\beta]=O(M^{\beta-1})$(见(1.177)—(1.178)), 故

$$\begin{aligned}(5.85)\text{的右端} &\leqslant (M-1)^{-2k\beta} C_1 M^{2\beta-2} \\ &\leqslant C_2 M^{-2k\beta+2\beta-2}. \end{aligned} \tag{5.86}$$

而

$$k-\frac{3}{4}>\frac{1}{2\beta}>\frac{1}{8}$$

故 $\beta<4$, 从而

$$2k\beta-2\beta+2>2\times\frac{7}{8}\beta-2\beta+2=-\frac{1}{4}\beta+2>1. \tag{5.87}$$

表明(5.86)所对应的左端的求和级数是收敛的:

$$\sum_{M=1}^{\infty} E\{(\sup_{[M^\beta]\leqslant N\leqslant[(M+1)^\beta]} |S(N)-N^{-k}[M^\beta]^k S([M^\beta])|)^2\}<+\infty, \tag{5.88}$$

即

$$\lim_{M\to\infty} \sup_{[M^\beta]\leqslant N\leqslant[(M+1)^\beta]} |S(N)-N^{-k}[M^\beta]^k S([M^\beta])|=0, \quad \text{a.s.}, \tag{5.89}$$

而

$$\begin{aligned} S(N) &\leqslant |S(N)-N^{-k}[M^\beta]^k S([M^\beta])|+N^{-k}[M^\beta]^k|S([M^\beta])| \\ &\leqslant \sup_{[M^\beta]\leqslant N\leqslant[(M+1)^\beta]} |S(N)-N^{-k}[M^\beta]^k S([M^\beta])| \\ &\quad +N^{-k}[M^\beta]^k|S([M^\beta])| \end{aligned} \tag{5.90}$$

$$<\frac{\varepsilon}{2}+\frac{\varepsilon}{2} \quad (\text{当}N\text{充分大}) \quad \text{a.s.}. \tag{5.91}$$

由(5.77), 表明

$$\sup_\lambda \left|\sum_{n=1}^{N} \xi_n e^{in\lambda}\right|=o(N^k), \quad \text{a.s.}, \quad k>\frac{7}{8}. \quad \blacksquare \tag{5.92}$$

以上证明的(5.92)是对弱P类平稳序列的有限 Fourier 变换在大样本条件下阶的估计。

特别要指出的是：上述结果对线性滤波过程也是成立的：设ξ_t为四阶实平稳列，η_t是线性滤波过程：

$$\eta_t = \sum_{k=-\infty}^{\infty} a_k \xi_{t-k}, \tag{5.93}$$

其中

(1) ξ_t是弱(或强)P序列；

(2) $$\sum_{k=-\infty}^{\infty} |a_k| < +\infty. \tag{5.94}$$

则η_t也是弱(或相应为强)P序列。

[64]得到了如下关于周期图的重对数律：

定理5.6 设ξ_n是线性过程

$$\xi_n = \sum_{k=0}^{\infty} c_k \varepsilon_{n-k} \tag{5.95}$$

满足

$$\sum_{k=0}^{\infty} k|c_k| < +\infty \tag{5.96}$$

和下列条件

(1) $\{\varepsilon_n\}$是独立同分布序列①；

(2) 存在一随机变量$Y(EY^2<+\infty)$和常数K，使得对$\forall x>0$，n

$$\Pr\{|\varepsilon(n)|>x\} \leqslant K\Pr\{|Y|>x\}, \tag{5.97}$$

则

$$\limsup_{N\to\infty} \frac{I_N(\lambda)}{\log\log N} = \begin{cases} f_\xi(\lambda), & \lambda \neq 0, \pm\pi, \\ 2f_\xi(\lambda), & \lambda = 0, \pm\pi, \end{cases} \quad \text{a.s}, \tag{5.98}$$

其中$f_\xi(\lambda)$是ξ_t的谱密度函数。

---

① 事实上还有鞅差等其它一组条件下也有类似的结果(见[64])。

[12],[65]都在各自假定的条件下得到了关于周期图的大样本性质。

3° 线性模型下周期图的大样本性质

下面讨论一种在应用中具有广泛用途的线性模型：

$$x_n = \sum_{k=1}^{P} \eta_k e^{in\lambda_k} + \xi_n, \tag{5.99}$$

其中

(1) ξ_n是满足弱P条件的实平稳列，$f_\xi(\lambda) \in L^2(d\lambda)$；

(2) $\{\eta_k\}_1^P$ 为随机变量，(δ_k, b_k)为常数，

$$0 < \delta_k \leqslant |\eta_k| \leqslant b_k < +\infty, \quad a.s.; \tag{5.100}$$

(3) $-\pi < \lambda_1 < \lambda_2 < \cdots < \lambda_P < \pi$，$P$ 为常数。 (5.101)

以下定理指出了在线性模型(5.99)下 x_n 的周期图的统计性质，特别在$\{\lambda_k\}_1^P$诸频率点上的性质。

定理5.7 设 x_n 是具(5.99)形式的线性模型。令

$$J_{xN}(\lambda) = N^{-\frac{31}{16}} \left| \sum_{n=1}^{N} x_n e^{in\lambda} \right|^2, \quad -\pi \leqslant \lambda \leqslant \pi, \tag{5.102}$$

则概率为1地当N充分大以后，存在与N无关的常数 c_0, c_1 使得

(1) $\inf\limits_{|\lambda-\lambda_j| \leqslant \frac{\pi}{N}} \{J_{xN}(\lambda)\} \geqslant c_0 N^{\frac{1}{16}}, \quad j = 1, 2, \cdots, P.$ (5.103)

(2) $\sup\limits_{\lambda \in \mathscr{A}} \{J_{xN}(\lambda)\} \leqslant c_1 N^{-\frac{1}{16}},$ (5.104)

其中

$$\mathscr{A} = \bigcap_{j=1}^{P} \{\lambda: N^{-\frac{15}{16}} \leqslant |\lambda - \lambda_j| \leqslant 2\pi - N^{-\frac{15}{16}}\}. \tag{5.105}$$

证明 以下只对 $P \neq 1$ 证明($P = 1$类似)。约定

$$S_N(\lambda; *_n) \triangleq \sum_{n=1}^{N} *_n e^{in\lambda}. \tag{5.106}$$

由(5.13)知

$$\inf_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}} |S_N(\lambda; e^{i*\lambda_j})|^2 = \inf_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}} \left(\frac{\sin\frac{N}{2}(\lambda-\lambda_j)}{\sin\frac{1}{2}(\lambda-\lambda_j)}\right)^2$$

$$= |D_N(\lambda-\lambda_j)|^2. \qquad (5.107)$$

当$|\lambda-\lambda_j|\leqslant\frac{\pi}{N}<\frac{2}{N}\pi$（第一个零点见图5.3)时$|D_N(\lambda-\lambda_j)|^2$的极

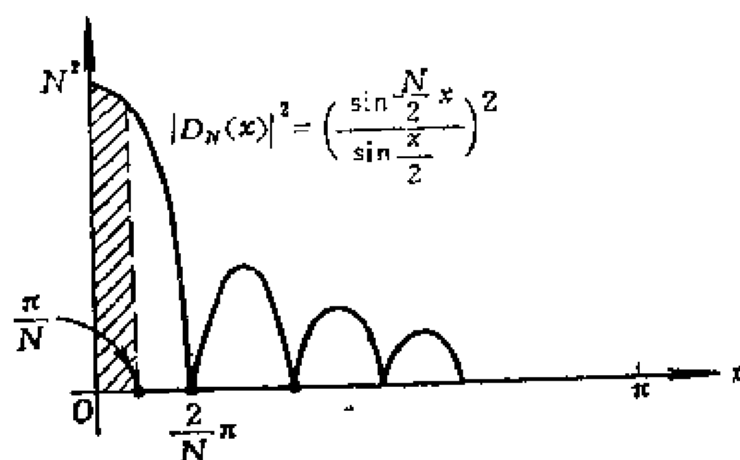


图 5.3

小值为

$$\left|D_N\left(\frac{\pi}{N}\right)\right|^2 = \frac{1}{\sin^2\frac{\pi}{2N}} \geqslant \frac{N^2}{\pi^2},$$

故

$$\inf_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}} |S_N(\lambda; e^{i*\lambda_j})|^2 \geqslant \frac{N^2}{\pi^2}. \qquad (5.108)$$

记

$$\lambda_0 = \lambda_p - 2\pi, \quad \delta_0 = \frac{1}{4}\min_{1\leqslant j\leqslant p}\{\lambda_j - \lambda_{j-1}\}. \qquad (5.109)$$

当N充分大，任取$\lambda\in\left[\lambda_j-\frac{\pi}{N},\ \lambda_j+\frac{\pi}{N}\right]$和$\lambda_k, k\neq j$，则

$$|\lambda-\lambda_k|\geqslant|\lambda_k-\lambda_j|-|\lambda-\lambda_j|\geqslant 4\delta_0-\frac{\pi}{N}\geqslant 2\delta_0,$$

$$(5.110)$$

$$|\lambda-\lambda_k|\leqslant|\lambda_k-\lambda_j|+|\lambda-\lambda_j|\leqslant|\lambda_p-\lambda_1|+\frac{\pi}{N}$$

$$\leqslant 2\pi-4\delta_0+\frac{\pi}{N}\leqslant 2\pi-2\delta_0. \qquad (5.111)$$

(5.111)最后第二个不等式利用了 $\lambda_p-\lambda_1=2\pi-(\lambda_1-\lambda_0)$。因此，当 $k\neq j$ 时

$$\sup_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}}|S_N(\lambda;e^{in\lambda_k})|=\sup_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}}\left|\frac{\sin\frac{N}{2}(\lambda-\lambda_k)}{\sin\frac{1}{2}(\lambda-\lambda_k)}\right|$$

$$\leqslant\sup_{|\lambda-\lambda_j|\leqslant\frac{\pi}{N}}\left\{\frac{1}{\left|\sin\frac{1}{2}(\lambda-\lambda_k)\right|}\right\}\leqslant\frac{1}{\sin\delta_0}. \qquad (5.112)$$

最后一个不等式利用了(5.110)和(5.111)。

$$|S_N(\lambda;x_n)|^2=\left|\sum_{n=1}^{N}\left[\sum_{j=1}^{p}\eta_j e^{in\lambda_j}+\xi_n\right]e^{in\lambda}\right|^2$$

$$=\left|\sum_{j=1}^{p}\eta_j S_N(\lambda;e^{in\lambda_j})+S_N(\lambda;\xi_n)\right|^2$$

$$=\left|\sum_{j=1}^{p}\eta_j S_N(\lambda;e^{in\lambda_j})\right|^2$$

$$+2\mathrm{Re}\left\{\sum_{j=1}^{p}\eta_j S_N(\lambda;e^{in\lambda_j})\overline{S_N(\lambda;\xi_n)}\right\}$$

$$+|S_N(\lambda;\xi_n)|^2 \qquad (5.113)$$

$$= |\eta_j|^2 |S_N(\lambda; e^{in\lambda_j})|^2$$

$$+ 2\mathrm{Re}\left\{\eta_j S_N(\lambda; e^{in\lambda_j}) \sum_{k \neq j} \overline{\eta_k} \overline{S_N}(\lambda; e^{in\lambda_k})\right\}$$

$$+ \left|\sum_{k \neq j} \eta_k S_N(\lambda; e^{in\lambda_k})\right|^2 + |S_N(\lambda; \xi_n)|^2$$

$$+ 2\mathrm{Re}\{\eta_j S_N(\lambda; e^{in\lambda_j}) \overline{S_N(\lambda; \xi_n)}\}$$

$$+ 2\mathrm{Re}\left\{\sum_{k \neq j} \eta_j S_N(\lambda; e^{in\lambda_k}) \overline{S_N(\lambda; \xi_n)}\right\}. \tag{5.114}$$

利用 $\mathrm{Re}(z) \geqslant -|z|$ 并将(5.114)的第三、四项舍去并用(5.100)的条件、(5.108)及(5.112)的结果，得

$$|S_N(\lambda; x_N)|^2$$

$$\geqslant |\delta_j|^2 \frac{N^2}{\pi^2} - 2|b_j| \frac{1}{\sin\delta_0} \sum_{l=1}^{p} |b_l| N$$

$$- 2|b_j| \frac{1}{\sin\delta_0} |S_N(\lambda; \xi_n)| - 2N|S_N(\lambda; \xi_n)| \sum_{k=1}^{p} |b_k|. \tag{5.115}$$

再由定理5.5的(5.71)式及(5.102)，当 N 充分大时

$$\inf_{|\lambda - \lambda_j| < \frac{\pi}{N}} \{J_{xN}(\lambda)\}$$

$$= N^{-\frac{31}{16}} \inf_{|\lambda - \lambda_j| < \frac{\pi}{N}} \{|S_N(\lambda; x_n)|^2\}$$

$$\geqslant N^{-\frac{31}{16}} \frac{|\delta_j|^2}{\pi^2} \left(N^2 - c_1 N - c_2 N^{\frac{7.5}{8}} \left(N^{-\frac{7.5}{8}} |S_N(\lambda; \xi_n)|\right)\right.$$

$$\left. - c_3 N^{1 + \frac{7.5}{8}} \left(N^{-\frac{7.5}{8}} |S_N(\lambda; \xi_n)|\right)\right)$$

$$\geqslant N^{-\frac{31}{16}} c_4 \left(N^2 - c_1 N - c_2 N^{\frac{7.5}{8}} - c_3 N^{1 + \frac{7.5}{8}}\right)$$

$$\geqslant N^{-\frac{31}{16}} c_4\left(N^2 - c_5 N^{1+\frac{7.5}{8}}\right) \geqslant c_0 N^{\frac{1}{16}} \quad \text{a.s.}. \tag{5.116}$$

由此证实了本定理的(5.103)。

再则，若 $\lambda \in \mathscr{A}$，由 $\mathscr{A}$ 的定义(5.105)

$$\frac{1}{2} N^{-\frac{15}{16}} \leqslant \frac{1}{2} |\lambda - \lambda_j| \leqslant \pi - \frac{1}{2} N^{-\frac{15}{16}},$$

$$|S_N(\lambda; e^{i \cdot \lambda_j})| = \left| \frac{\sin \frac{N}{2}(\lambda - \lambda_j)}{\sin \frac{1}{2}(\lambda - \lambda_j)} \right| \leqslant \frac{1}{\left|\sin \frac{1}{2} N^{-\frac{15}{16}}\right|}$$

$$\leqslant C N^{\frac{15}{16}}. \tag{5.117}$$

利用 $\mathrm{Re}(z) \leqslant |z|$，则由(5.114)，(5.117)，(5.92)

$$\begin{aligned}
&\sup_{\lambda \in \mathscr{A}} |S_N(\lambda; x_n)|^2 \\
&\leqslant C |b_j|^2 N^{\frac{15}{16} \times 2} + 2 |b_j| C N^{\frac{15}{16}} \sum_{k \neq j} |b_k| C N^{\frac{15}{16}} \\
&\quad + \left(\sum_{k \neq j} |b_k| C N^{\frac{15}{16}}\right)^2 + \left(c_2 N^{\frac{15}{16}}\right)^2 \\
&\quad + 2 |b_j| C N^{\frac{15}{16}} c_1 N^{\frac{15}{16}} \\
&\quad + 2 \sum_{k \neq j} |b_k| C N^{\frac{15}{16}} c_1 N^{\frac{15}{16}} \\
&\leqslant c_3 N^{\frac{30}{16}} \quad \text{a.s.}.
\end{aligned} \tag{5.118}$$

于是

$$\sup_{\lambda \in \mathscr{A}} \left(N^{-\frac{31}{16}} |S_N(\lambda; x_n)|^2\right)$$

$$= \sup_{\lambda \in \mathscr{A}} (J_{xN}(\lambda)) \leqslant c_4 N^{-\frac{31}{16}} N^{\frac{30}{16}}$$

$$= c_4 N^{-\frac{1}{16}} \quad \text{a.s.}. \quad \blacksquare \tag{5.119}$$

定理5.7的重要在于揭示如下事实：对于符合(5.99)的线性模型（显见，x_n的谱函数并非绝对连续的），x_n的周期图$I_N(\lambda)$当λ靠近$\{\lambda_j\}$时

$$I_N(\lambda) = \frac{1}{2\pi N} |S_N(\lambda, x_n)|^2 = \frac{N^{\frac{31}{61}}}{2\pi N} \{J_{Nx}(\lambda)\}$$

$$\geqslant C N^{\frac{15}{16}} N^{\frac{1}{16}} = CN, \tag{5.120}$$

即$I_N(\lambda)$在$\{\lambda_j\}_1^p$诸频率附近至少有N的增长量，而稍离开这些离散的频率点，则至多有

$$\sup_{\lambda \in \mathscr{A}} \{I_N(\lambda)\} \leqslant c_1 N^{\frac{15}{16}} N^{-\frac{1}{16}} = c_1 N^{\frac{14}{16}} \tag{5.121}$$

的增长量。因此，当N充分大时，可以想象$I_N(\lambda)$在诸$\{\lambda_k\}_1^p$有非常锐的尖峰。这就提供了从$\{x_n\}_1^N$观测样本中识别$\{\lambda_k\}_1^p$诸离散频率的理论依据。

特别需要指出的是：由(5.100)的条件表明线性模型包括了如下在工程技术中极具典型的包含确定型周期趋势项加平稳扰动的模型

$$x_n = \sum_{k=1}^{p} A_k e^{in\lambda_k} + \xi_n, \tag{5.122}$$

$$x_n = \sum_{k=1}^{p} A_k \cos(\lambda_k n + \psi_k) + \xi_n. \tag{5.123}$$

§3 加窗谱密度估计方法

在上一节中已证明了，当x_t是有谱密度时，周期图$I_N(\lambda)$虽

为 $f_x(\lambda)$的渐近无偏估计量，但它不是相容估计，因此需要加以改造以便得到好的统计性质，其中一条途径就是“加窗”，即对周期图 $I_N(\lambda)$进行加权平均以减少它的方差。

1° 加窗谱估计及其统计性质

由概率论的知识知道：设$\{\xi_n\}$是独立同分布的随机变量，则 $\bar{\xi}_N = \frac{1}{N}\sum_{k=1}^{N}\xi_k$的方差仅为原来的$\frac{1}{N}$。根据(5.50)知，若对$I_N(\lambda)$的不同点的变量加以平均就有可能减少它的方差使它变成相容估计量。

考虑新的估计量

$$f_N(\lambda) = \int_{\Pi} W_N(\lambda-\mu) I_N(\mu) \mathrm{d}\mu, \qquad (5.124)$$

其中 $W_N(\mu)$是定义于 Π 上的谱窗函数：

(1) $W_N(\mu)$是一串定义于 Π 上的实函数，满足

$$\int_{\Pi} W_N(\mu)\mathrm{d}\mu = 1; \quad \int_{\Pi} W_N^2(\mu)\mathrm{d}\mu < +\infty. \qquad (5.125)$$

(2) 对任给的 $\varepsilon>0$，对$|\mu|\geqslant\varepsilon$ 均匀地有

$$W_N(\mu)\to 0 \quad (N\to\infty). \qquad (5.126)$$

(3) 对常数 $A>0$，当 $N\to\infty$，有

$$\max_{|\mu|<\frac{A}{N}}\left|\frac{\int_{\Pi} W_N(\lambda)W_N(\lambda+\mu)\mathrm{d}\lambda}{\int_{\Pi} W_N^2(\lambda)\mathrm{d}\lambda} - 1\right| \to 0. \qquad (5.127)$$

(4) $W_N(-\mu) = W_N(\mu), \quad \mu\in\Pi.$ (5.128)

以上表明：$W_N(\mu)$是一串实的、偶的、规范的函数序列，并且随着N的增大$\{W_N(\mu)\}$这一串函数其“质量”将愈来愈集中在原点 $\mu=0$ 附近，以后我们均假定 Π 以外按周期来开拓。对于(5.127)的作用在以下的证明中将看清楚。

以下定理说明用(5.124)定义的估计量是相容估计。

定理5.8 设 x_t 是正态平稳列，谱密度 $f_x(\lambda)$ 具有一级连续导函数，$W_N(\mu)$ 是满足(5.125)—(5.128)的谱窗函数列，则由(5.124)定义的估计量具有性质：

(1)
$$Ef_N(\lambda)=f_x(\lambda)+\int_\Pi W_N(\lambda-\mu)[f_x(\mu)-f_x(\lambda)]\mathrm{d}\mu+O\left(\frac{\log N}{N}\right),\quad -\pi\leqslant\lambda\leqslant\pi. \tag{5.129}$$

(2)
$$D(f_N(\lambda))\sim\frac{2\pi}{N}\left(\int_\Pi W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha+\int_\Pi W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha\right)^{①}. \tag{5.130}$$

证明 先证(5.129)。由(5.124)，显见

$$E(f_N(\lambda))=\int_\Pi W_N(\lambda-\mu)EI_N(\mu)\mathrm{d}\mu$$
$$=\int_\Pi W_N(\lambda-\mu)\left(\int_\Pi\frac{1}{2\pi N}\frac{\sin^2\frac{N}{2}(\mu-\alpha)}{\sin^2\frac{1}{2}(\mu-\alpha)}f_x(\alpha)\mathrm{d}\alpha\right)\mathrm{d}\mu. \tag{5.131}$$

利用(5.15)及(5.125)可得

$$E(f_N(\lambda))=f_x(\lambda)+\int_\Pi W_N(\lambda-\mu)\int_\Pi\frac{1}{2\pi N}\frac{\sin^2\frac{N}{2}(\mu-\alpha)}{\sin^2\frac{1}{2}(\mu-\alpha)}$$

---

① $D(f_N(\lambda))$ 与(5.130)差一尾项 e_N，

$|e_N|\leqslant C\log N/N+\varepsilon_N\left(\frac{1}{N}\int_\Pi W_N^2(x)\,\mathrm{d}x\right)$，$\varepsilon_N\to0$，$N\to\infty$。

$$\times[f_x(\alpha)-f_x(\lambda)]\mathrm{d}\alpha\mathrm{d}\mu. \tag{5.132}$$

而在 $f_x(\lambda)$ 有一级连续偏导数之下，由定理 5.2 证明中的(5.16)—(5.17)知，(5.132)右端的第二重积分可用(5.12)来估计，再由其阶对变量的一致性可得

$$\begin{aligned} b_N(\lambda) &= E(f_N(\lambda)) - f_x(\lambda) \\ &= \int_\Pi W_N(\lambda-\mu)\left\{(f_x(\mu)-f_x(\lambda)) + O\left(\frac{\log N}{N}\right)\right\}\mathrm{d}\mu \\ &= \int_\Pi W_N(\lambda-\mu)(f_x(\mu)-f_x(\lambda))\mathrm{d}\mu + O\left(\frac{\log N}{N}\right). \end{aligned} \tag{5.133}$$

以下来证第二个估计式。由(5.29)显见

$$\begin{aligned} D(f_N(\lambda)) &= \int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)\mathrm{Cov}(I_N(\alpha),I_N(\beta))\mathrm{d}\alpha\mathrm{d}\beta \\ &= \int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)\left[\frac{f_x(\alpha)f_x(\beta)}{N^2}\right. \\ &\quad \left.\times\{D_N^2(\alpha-\beta)+D_N^2(\alpha+\beta)\} + O\left(\frac{\log N}{N}\right)\right]\mathrm{d}\alpha\mathrm{d}\beta \\ &= \frac{1}{N^2}\int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)D_N^2(\alpha-\beta)f_x(\alpha) \\ &\quad \times f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta \\ &\quad + \frac{1}{N^2}\int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)D_N^2(\alpha+\beta) \\ &\quad \times f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta + O\left(\frac{\log N}{N}\right) \\ &\triangleq I_1 + I_2 + O\left(\frac{\log N}{N}\right). \end{aligned} \tag{5.134}$$

由于 I_1 和 I_2 类似，以下只对 I_1 估计。利用函数的周期性：

$$I_1 = \frac{1}{N^2}\int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)D_N^2(\alpha-\beta)f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta$$

$$=\frac{1}{N^2}\left\{\int_{|u|<\frac{A}{N}}\int_{\Pi}D_N^2(u)W_N(\lambda-\alpha)W_N(\lambda+u-\alpha)f_x(\alpha)\right.$$

$$\times f_x(\alpha-u)\mathrm{d}\alpha\mathrm{d}u+\int_{|u|>\frac{A}{N}}\int_{\Pi}D_N^2(u)W_N(\lambda-\alpha)W_N(\lambda+u-\alpha)$$

$$\left.\times f_x(\alpha)f_x(\alpha-u)\mathrm{d}\alpha\mathrm{d}u\right\}\triangleq A_1+A_2. \tag{5.135}$$

$$A_1=\frac{1}{N^2}\int_{|u|<\frac{A}{N}}D_N^2(u)\left\{\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda+u-\alpha)\right.$$

$$\left.\times f_x(\alpha)f_x(\alpha-u)\mathrm{d}\alpha\right\}\mathrm{d}u. \tag{5.136}$$

但(5.136)中括号部分可用 $\int_{\Pi}W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha$ 估计,因为:

$$\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)f_x(\alpha)f_x(\alpha-u)\mathrm{d}\alpha$$

$$-\int_{\Pi}W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha$$

$$=\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)f_x(\alpha)$$

$$\times[f_x(\alpha)-uf_x'(\bar{\theta})]\mathrm{d}\alpha-\int_{\Pi}W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha$$

$$=\int_{\Pi}(W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)-W_N^2(\lambda-\alpha))f_x^2(\alpha)\mathrm{d}\alpha$$

$$-u\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)f_x(\alpha)f_x'(\bar{\theta})\mathrm{d}\alpha$$

$$\triangleq J_1+J_2 \tag{5.137}$$

(其中 $\bar{\theta}$ 位于 α 和 u 的中间值)。以下分别估 J_1 和 J_2。

$$|J_2|\leqslant|u|\int_{\Pi}|f_x(\alpha)||f'_x(\bar{\theta})||W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)|\mathrm{d}\alpha$$

$$\leqslant|u|(|f'||f|)_{\max}\int_{\Pi}|W_N(\lambda-\alpha)||W_N(\lambda-\alpha+u)|\mathrm{d}\alpha$$

$$\leqslant|u|(|f'||f|)_{\max}\left\{\int_{\Pi}W_N^2(\lambda-\alpha)\mathrm{d}\alpha\int_{\Pi}W_N^2(\lambda-\alpha+u)\mathrm{d}\alpha\right\}^{\frac{1}{2}}$$

$$\leqslant\frac{C}{N}\int_{\Pi}W_N^2(\lambda)\mathrm{d}\lambda=o\left(\int_{\Pi}W_N^2(\lambda)\mathrm{d}\lambda\right),\quad N\to\infty,\qquad(5.138)$$

其中(5.138)最后一个不等式利用了$W_N(\lambda)$的周期性和偶性及$|u|\leqslant\frac{A}{N}$。不难看出(5.138)对λ是一致的。

$$|J_1|\leqslant|f|_{\max}^2\int_{\Pi}|W_N(\lambda-\alpha)W_N(\lambda-\alpha+u)-W_N^2(\lambda-\alpha)|\mathrm{d}\alpha$$

$$\leqslant|f|_{\max}^2\left\{\int_{\Pi}W_N^2(\mu)\mathrm{d}\mu\right\}^{\frac{1}{2}}\left\{\int_{\Pi}(W_N(\mu-u)-W_N(\mu))^2\mathrm{d}\mu\right\}^{\frac{1}{2}}$$

$$=|f|_{\max}^2\left\{\int_{\Pi}W_N^2(\mu)\mathrm{d}\mu\right\}^{\frac{1}{2}}\left\{\int_{\Pi}[W_N^2(\mu-u)+W_N^2(\mu)\right.$$

$$\left.-2W_N(\mu-u)W_N(\mu)]\mathrm{d}\mu\right\}^{\frac{1}{2}}$$

$$=\sqrt{2}\,|f|_{\max}^2\left\{\int_{\Pi}W_N^2(\mu)\mathrm{d}\mu\right\}^{\frac{1}{2}}$$

$$\times\left\{\int_{\Pi}(W_N^2(\mu)-W_N(\mu-u)W_N(\mu))\mathrm{d}\mu\right\}^{\frac{1}{2}}.\qquad(5.139)$$

然而由$W_N(\lambda)$的性质(5.127)，当$|u|\leqslant\frac{A}{N}$时，N充分大之下有

$$\left|\frac{\int_{\Pi} W_N(\mu) W_N(\mu+u)\,\mathrm{d}\mu}{\int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu}-1\right| = o(1), \qquad (5.140)$$

或

$$\left|\int_{\Pi} W_N(\mu) W_N(\mu+u)\,\mathrm{d}\mu - \int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu\right|$$

$$= o\left(\int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu\right). \qquad (5.141)$$

注意(5.140)对$|u|\leqslant A/N$是一致地成立。将(5.141)代入(5.139)得(对$|u|\leqslant A/N$也是一致的)：

$$|J_1| = o\left(\int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu\right). \qquad (5.142)$$

结合(5.138)A_1可改写为：

$$A_1 = \frac{1}{N^2}\int_{|u|\leqslant \frac{A}{N}} D_N^2(u)\left\{\int_{\Pi} W_N^2(\lambda-\alpha) f_x^2(\alpha)\,\mathrm{d}\alpha\right.$$

$$\left. + o\left(\int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu\right)\right\}\mathrm{d}u. \qquad (5.143)$$

注意到

$$\lim_{N\to\infty} \frac{1}{2\pi N}\int_{|u|\leqslant \varepsilon} D_N^2(u)\,\mathrm{d}u = 1, \qquad (5.144)$$

则

$$A_1 = \frac{2\pi}{N}\int_{\Pi} W_N^2(\lambda-\alpha) f_x^2(\alpha)\,\mathrm{d}\alpha + o\left(\frac{1}{N}\int_{\Pi} W_N^2(\mu)\,\mathrm{d}\mu\right). \quad (5.145)$$

以下再来估计A_2。

$$A_2 = \frac{1}{N^2}\int_{|u|>\frac{A}{N}}\int_{\Pi} D_N^2(u) W_N(\lambda-\alpha) W_N(\lambda+u-\alpha) f_x(\alpha)$$

$$\times f_x(a-u)\,\mathrm{d}a\,\mathrm{d}u.$$

$$|A_2| \leqslant \frac{1}{N^2}\int_{|u|>\frac{A}{N}} D_N^2(u)\left|\int_{\Pi} W_N(\lambda-a)W_N(\lambda+u-a)\right.$$

$$\left.\times f_x(a)f_x(a-u)\,\mathrm{d}a\right|\mathrm{d}u$$

$$\leqslant \frac{|f|_{\max}^2}{N^2}\int_{|u|>\frac{A}{N}} D_N^2(u)\left(\int_{\Pi}|W_N(x)W_N(x-u)|\,\mathrm{d}x\right)\mathrm{d}u$$

$$\leqslant \frac{c_1}{N^2}\int_{|u|>\frac{A}{N}} \frac{\mathrm{d}u}{\sin^2\frac{u}{2}}\left(\int_{\Pi} W_N^2(x)\,\mathrm{d}x\right). \tag{5.146}$$

但是

$$\int_{|u|>\frac{A}{N}} \frac{\mathrm{d}u}{\sin^2\frac{u}{2}} = 2\operatorname{ctg}\frac{A}{2N} = 2\frac{\cos\frac{A}{2N}}{\sin\frac{A}{N}}$$

$$\leqslant 2\cdot\frac{2N}{A}\left(\frac{\sin\frac{A}{2N}}{\frac{A}{2N}}\right)^{-1} \leqslant C_2\frac{N}{A}, \tag{5.147}$$

其中(5.147)最后一个不等式是因为当N充分大之下

$$\left(\frac{\sin\frac{A}{2N}}{\frac{A}{2N}}\right) > \frac{1}{2}.$$

因此

$$|A_2| \leqslant \frac{C_3}{A}\left(\frac{1}{N}\int_{\Pi} W_N^2(x)\,\mathrm{d}x\right). \tag{5.148}$$

由于C_1, C_2, C_3都不依赖于A，以及A选择的随意性，所以

$$A_2 = o\left(\frac{1}{N}\int_{\Pi} W_N^2(x)\,\mathrm{d}x\right). \tag{5.149}$$

因此

$$I_1 = \frac{2\pi}{N}\int_\Pi W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha + o\left(\frac{1}{N}\int_\Pi W_N^2(x)\mathrm{d}x\right). \quad (5.150)$$

至于(5.134)中的I_2:

$$I_2 = \frac{1}{N^2}\int_\Pi\int_\Pi W_N(\lambda-\alpha)W_N(\lambda-\beta)D_N^2(\alpha+\beta)f_x(\alpha)f_x(\beta)\mathrm{d}\alpha\mathrm{d}\beta$$

可以用

$$\frac{2\pi}{N}\int_\Pi W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha \quad (5.151)$$

来估计，其方式和I_1完全类似，此处从略。综上所述，可得：

$$\begin{aligned} &\mathrm{D}(f_N(\lambda)) \\ &= \frac{2\pi}{N}\int_\Pi W_N^2(\lambda-\alpha)f_x^2(\alpha)\mathrm{d}\alpha \\ &+ \frac{2\pi}{N}\int_\Pi W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha \\ &+ o\left(\frac{1}{N}\int_\Pi W_N^2(x)\mathrm{d}x\right) + O\left(\frac{\log N}{N}\right). \quad \blacksquare \end{aligned} \quad (5.152)$$

推论 1 在定理5.8的条件下

(1) $\lim\limits_{N\to\infty} Ef_N(\lambda) = f_x(\lambda)$.

(2)

$$\mathrm{D}(f_N(\lambda)) \sim \begin{cases} \frac{2\pi}{N}f_x^2(\lambda)\int_\Pi W_N^2(x)\mathrm{d}x, & \lambda \neq 0, \pm\pi; \\ \frac{4\pi}{N}f_x^2(\lambda)\int_\Pi W_N^2(x)\mathrm{d}x, & \lambda = 0, \pm\pi. \end{cases}$$

(5.153)

证明 以下只对W_N是非负的场合给出证明。由(5.133)及(5.126)当N充分大时，

$$|b_N(\lambda)| \leqslant \left|\int_{|\lambda-\mu|<\varepsilon} W_N(\lambda-\mu)f_x'(\bar{\theta})(\mu-\lambda)\mathrm{d}\mu\right|$$

$$+\left|\int_{|\lambda-\mu|>\varepsilon} W_N(\lambda-\mu)(f_x(\lambda)-f_x(\mu))\mathrm{d}\mu\right|$$

$$+K\frac{\log N}{N}\triangleq I_1+I_2+I_3. \tag{5.154}$$

$$|I_1|\leqslant|f'|_{\max}\varepsilon\int_{|\lambda-\mu|<\varepsilon} W_N(\lambda-\mu)\mathrm{d}\mu<c_1\varepsilon. \tag{5.155}$$

又，当N充分大时，由(5.126)可得

$$|I_2|\leqslant 2|f|_{\max}\int_{|\lambda-\mu|>\varepsilon} W_N(\lambda-\mu)\mathrm{d}\mu<c_2\varepsilon. \tag{5.156}$$

而I_3当N充分大时可任意小，故$b_N(\lambda)\to 0$，由此得(1)。

以下再证(5.153)。当$\lambda\neq 0,\pm\pi$时，由(5.152)

$$D(f_N(\lambda))=\frac{2\pi}{N}\left[\int_{|\lambda-\mu|<\varepsilon} W_N^2(\lambda-\mu)(f_x^2(\lambda)\right.$$

$$+2(\lambda-\mu)f_x(\bar\theta)f_x'(\bar\theta))\mathrm{d}\mu$$

$$\left.+\int_{|\lambda-\mu|\geqslant\varepsilon} W_N^2(\lambda-\mu)f_x^2(\mu)\mathrm{d}\mu\right]$$

$$+\frac{2\pi}{N}\int_\Pi W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha$$

$$+o\left(\frac{1}{N}\int_\Pi W_N^2(x)\mathrm{d}x\right)+O\left(\frac{\log N}{N}\right).$$

故

$$\left|D(f_N(\lambda))-\frac{2\pi}{N}f_x^2(\lambda)\int_\Pi W_N^2(x)\mathrm{d}x\right|$$

$$\leqslant\frac{2\pi}{N}\left[\int_{|\lambda-\mu|<\varepsilon} 2W_N^2(\lambda-\mu)|\lambda-\mu||f_x(\bar\theta)||f_x'(\bar\theta)|\mathrm{d}\mu\right]$$

$$+\frac{2\pi}{N}\int_{|\lambda-\mu|>\varepsilon} W_N^2(\lambda-\mu)|f_x^2(\mu)-f_x^2(\lambda)|\mathrm{d}\mu$$

$$+\frac{2\pi}{N}\left|\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha\right|$$

$$+\left(o\left(\frac{1}{N}\int_{\Pi}W_N^2(x)\mathrm{d}x\right)+O\left(\frac{\log N}{N}\right)\right)$$

$$\triangleq J_1+J_2+J_3+\theta_N. \tag{5.157}$$

$$\left.\begin{aligned}|J_1|&\leqslant\varepsilon\frac{c_1}{N}\int_{|\lambda-\mu|<\varepsilon}W_N^2(\lambda-\mu)\mathrm{d}\mu,\\|J_2|&\leqslant\frac{c_2}{N}\int_{|\lambda-\mu|\geqslant\varepsilon}W_N^2(\lambda-\mu)\mathrm{d}\mu\leqslant\frac{K}{N}\varepsilon.\end{aligned}\right\} \tag{5.158}$$

$|J_2|$的最后一个不等式是：对$|\theta|\geqslant\varepsilon$，

$$W_N^2(\theta)\to0(\text{均匀})\text{当}N\to\infty.$$

$$J_3=\frac{2\pi}{N}\int_{\Pi}W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha$$

$$=\frac{2\pi}{N}\left(\int_{|\lambda-\alpha|<\varepsilon}W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha\right.$$

$$\left.+\int_{|\lambda-\alpha|\geqslant\varepsilon}W_N(\lambda-\alpha)W_N(\lambda+\alpha)f_x^2(\alpha)\mathrm{d}\alpha\right)$$

$$=\frac{2\pi}{N}(J_{3,1}+J_{3,2}), \tag{5.159}$$

其中

$$|J_{3,1}|\leqslant|f^2|_{\max}\int_{|\lambda-\alpha|<\varepsilon}|W_N(\lambda-\alpha)W_N(\lambda+\alpha)|\mathrm{d}\alpha$$

$$\leqslant|f^2|_{\max}\left\{\int_{|\lambda-\alpha|<\varepsilon}W_N^2(\lambda-\alpha)\mathrm{d}\alpha\right\}^{\frac{1}{2}}\left\{\int_{|\lambda-\alpha|<\varepsilon}W_N^2(\lambda+\alpha)\mathrm{d}\alpha\right\}^{\frac{1}{2}}.$$

当ε取适当小时，$|\lambda+\alpha|\geqslant\varepsilon$（当$\alpha$满足$|\lambda-\alpha|<\varepsilon,\lambda\neq0$）。故当$N$充分大时

$$\int_{|\lambda-\alpha|<\varepsilon}W_N^2(\lambda+\alpha)\mathrm{d}\alpha\to0, \tag{5.160}$$

$|J_{3,1}|\leqslant\epsilon\left(\int_{\Pi}W_N^2(x)\mathrm{d}x\right)$。类似方法可得$|J_{3,2}|\leqslant\epsilon\left(\int_{\Pi}W_N^2(x)\mathrm{d}x\right)$，$\epsilon$任意小。因此

$$D(f_N(\lambda))=\frac{2\pi}{N}f_x(\lambda)\int_{\Pi}W_N^2(x)\mathrm{d}x+o\left(\frac{1}{N}\int_{\Pi}W_N^2(x)\mathrm{d}x\right)+O\left(\frac{\log N}{N}\right),$$

或

$$D(f_N(\lambda))\sim\frac{2\pi}{N}f_x(\lambda)\int_{\Pi}W_N^2(x)\mathrm{d}x。\tag{5.161}$$

又，当$\lambda=0,\pm\pi$时，可利用$W_N(x)$的偶性和周期性，比如$\lambda=0$，(5.152)变成为：

$$D(f_N(0))=\frac{4\pi}{N}\int_{\Pi}W_N^2(\alpha)f_x^2(\alpha)\mathrm{d}\alpha+o\left(\frac{1}{N}\int_{\Pi}W_N^2(x)\mathrm{d}x\right)+O\left(\frac{\log N}{N}\right)。\tag{5.162}$$

再用以上类似的方法可估得

$$D(f_N(0))\sim\frac{4\pi}{N}f_x^2(0)\int_{\Pi}W_N^2(\alpha)\mathrm{d}\alpha。\ \blacksquare\tag{5.163}$$

2° 谱窗函数的选择

以上介绍了加窗谱密度估计$f_N(\lambda)$是渐近无偏的估计，并且由定理5.8的推论知，只要选$W_N(x)$使得

$$\int_{\Pi}W_N^2(x)\mathrm{d}x$$

的增长比N来得慢时，由(5.161)和(5.163)可看出$f_N(\lambda)$将是相容估计量，$D(f_N(\lambda))\to0$（当$N\to\infty$），并由此知它也是L^2收敛的

$$f_N(\lambda)\xrightarrow{L^2}f(\lambda)。\tag{5.164}$$

如何选择$W_N(\lambda)$？通常方法是：

(a) 选一核函数$K(t)$，它是实的、偶的并且除个别点以外皆可导的函数，且满足

$$\begin{cases} K(0)=1, & (5.165) \\ K(t)=0, \quad |t|>1. & (5.166) \end{cases}$$

(b) 选一串正整数列$\{m_N\}$满足

(i) $m_N \to +\infty$; (5.167)

(ii) $m_N N^{-1} \to 0$. (5.168)

(c) 令 $w_N(k)=K\left(\frac{k}{m_N}\right)$, $k=0,\pm 1,\pm 2,\cdots$, (5.169)

$$\begin{aligned} W_N(\lambda) &= \frac{1}{2\pi}\sum_{k=-m_N}^{m_N} K\left(\frac{k}{m_N}\right)e^{-ik\lambda} \\ &= \frac{1}{2\pi}\sum_{k=-m_N}^{m_N} w_N(k)e^{-ik\lambda}, \quad \lambda\in\Pi. \end{aligned} \tag{5.170}$$

以后称$\{w_N(k)\}$为窗系数，$W_N(\lambda)$为谱窗函数，$K(t)$为谱窗的核函数，m_N为滞后值。

最简单的核函数为

$$K(t)=\begin{cases} 1, & |t|\leqslant 1; \\ 0, & |t|>1. \end{cases} \tag{5.171}$$

m_N 列选为$[\sqrt{N}]$，则(5.166)—(5.168)皆满足，此时

$$w_N(k)=\begin{cases} 1, & |k|\leqslant m_N; \\ 0, & \text{其它}, \end{cases} \tag{5.172}$$

并且

$$W_N(\lambda)=\frac{1}{2\pi}\sum_{k=-m_N}^{m_N} e^{-ik\lambda}=\frac{1}{2\pi}\,\frac{\sin\left(m_N+\frac{1}{2}\right)\lambda}{\sin\frac{1}{2}\lambda}. \tag{5.173}$$

在以上条件下确定的 $W_N(\lambda)$ 皆满足 (5.125)—(5.128)对谱窗函数的基本要求。

(1) 由(5.170)得

$$\int_{\Pi} W_N(\lambda)\mathrm{d}\lambda = \frac{1}{2\pi}\sum_{k=-m_N}^{m_N} K\left(\frac{k}{m_N}\right)\int_{\Pi} e^{-ik\lambda}\mathrm{d}\lambda$$

$$= \frac{1}{2\pi}\cdot K(0)\int_{\Pi}\mathrm{d}\lambda = 1.$$

(2) $$W_N(-\lambda) = \frac{1}{2\pi}\sum_{k=-m_N}^{m_N} K\left(\frac{k}{m_N}\right)e^{ik\lambda}$$

$$= \frac{1}{2\pi}\sum_{\mu=-m_N}^{m_N} K\left(-\frac{\mu}{m_N}\right)e^{-i\mu\lambda}$$

$$= \frac{1}{2\pi}\sum_{\mu=-m_N}^{m_N} K\left(\frac{\mu}{m_N}\right)e^{-i\mu\lambda} = W_N(\lambda).$$

(3) 若 $K(x)$ 还满足几乎处处可导的条件，则（见[66]）

$$F(t) = \int_{R_1} K(x)e^{-itx}\mathrm{d}x = o\left(\frac{1}{|t|}\right). \qquad (5.174)$$

从而

$$\frac{1}{2\pi}\int_{-1}^{1} K(x)e^{-i(m_N\lambda)x}\mathrm{d}x = o\left(\frac{1}{m_N\lambda}\right). \qquad (5.175)$$

当 $|\lambda|\geqslant\varepsilon$ 时

$$\frac{1}{2\pi}\int_{-1}^{1} K(x)e^{-i(m_N\lambda)x}\mathrm{d}x = o\left(\frac{1}{m_N}\right) \qquad (5.176)$$

一致地成立。因此当 $N\to\infty$ 时，

$$W_N(\lambda) = \frac{1}{2\pi}\sum_{|k|\leqslant m_N} K\left(\frac{k}{m_N}\right)e^{-ik\lambda}$$

$$= \frac{m_N}{2\pi}\left\{\int_{-1}^{1} K(x)e^{-ix(m_N\lambda)}\mathrm{d}x + \varepsilon_N\right\}, \qquad (5.177)$$

其中 ε_N 比括号中的积分值是高阶无穷小量。又由(5.176)，$N\to$

∞时，$W_N(\lambda)\to 0$，一致地对$|\lambda|\geqslant\varepsilon$成立。

(4) 再考查条件(5.127)。

$$\Delta_\mu \triangleq \left| \frac{\int_\Pi W_N(\lambda)W_N(\lambda+\mu)\mathrm{d}\lambda}{\int_\Pi W_N^2(\lambda)\mathrm{d}\lambda} - 1 \right|$$

$$= \left| \frac{\frac{1}{2\pi}\sum\limits_{|k|\leqslant m_N} K^2\left(\frac{k}{m_N}\right)\mathrm{e}^{-\mathrm{i}k\mu}}{\frac{1}{2\pi}\sum\limits_{|k|\leqslant m_N} K^2\left(\frac{k}{m_N}\right)} - 1 \right|$$

$$= \frac{\left| m_N\int_{-1}^{1} K^2(x)(\mathrm{e}^{-\mathrm{i}x(m_N\mu)}-1)\mathrm{d}x + o(m_N)\right|}{\left| m_N\int_{-1}^{1} K^2(x)\mathrm{d}x + o(m_N)\right|}.$$

(5.178)

当$\mu=0$时，$\Delta_\mu=0$；

当$|\mu|\leqslant\frac{A}{N}$时，$|m_N\mu|\leqslant A\frac{m_N}{N}<\varepsilon$（当$N$充分大）。

$$|\Delta_\mu| \leqslant \frac{\int_{-1}^{1} K^2(x)\,|\mathrm{e}^{-\mathrm{i}x(m_N\mu)}-1|\,\mathrm{d}x}{\|K\|^2+o(1)} + \varepsilon_N \qquad (5.179)$$

（其中$\varepsilon_N\to 0$，当$N\to\infty$），又

$$|\mathrm{e}^{-\mathrm{i}x(m_N\mu)}-1|^2 = [1+1-2\cos(xm_N\mu)]$$

$$= 2(1-\cos(xm_N\mu)) = 2\left(2\sin^2\frac{xm_N\mu}{2}\right),$$

即

$$|\mathrm{e}^{-\mathrm{i}x(m_N\mu)}-1| \leqslant 2\frac{|xm_N\mu|}{2} = |xm_N\mu| \leqslant |m_N\mu| < \varepsilon, \quad (5.180)$$

故代入(5.179)得

$$\max_{|\mu| \le \frac{A}{N}} |\Delta_\mu| \leqslant \frac{\varepsilon' \|K\|^2}{\|K\|^2} + \varepsilon_N = o(1).$$

表明(5.127)条件满足，其中$\|K\|^2 = \int_{-1}^{1} K^2(t)\mathrm{d}t$。

定理5.8推论2 在定理5.8条件下，

$$D(f_N(\lambda)) \sim \begin{cases} \frac{m_N}{N} f_*^2(\lambda)\|K\|^2, & \text{当 } \lambda \neq 0, \pm\pi; \\ 2\frac{m_N}{N} f_*^2(\lambda)\|K\|^2, & \text{当 } \lambda = 0, \pm\pi. \end{cases} \tag{5.181}$$

证明 由 Bessel 等式

$$\int_\Pi W_N^2(\lambda)\mathrm{d}\lambda = \frac{1}{2\pi}\sum_{|k|\le m_N}(w_N(k))^2$$

$$= \frac{m_N}{2\pi}\left(\sum_{|k|\le m_N} K^2\left(\frac{k}{m_N}\right)\frac{1}{m_N}\right)$$

$$\sim \frac{m_N}{2\pi}\int_{-1}^{1} K^2(t)\mathrm{d}t = \frac{m_N}{2\pi}\|K\|^2.$$

因此对$\lambda \neq 0, \pm\pi$，由定理5.8

$$\mathrm{D}(f_N(\lambda)) \sim \frac{2\pi}{N}\left\{\frac{m_N}{2\pi}\|K\|^2\right\} f_*^2(\lambda) = \frac{m_N}{N} f_*^2(\lambda)\|K\|^2. \tag{5.175'}$$

类似地对$\lambda = 0, \pm\pi$，

$$\mathrm{D}(f_N(\lambda)) \sim 2\frac{m_N}{N} f_*^2(\lambda)\|K\|^2. \quad \blacksquare \tag{5.176'}$$

定理5.8推论3 在定理5.8条件下，$f_N(\lambda)$是$f_x(\lambda)$的相合估计量，即当$N\to\infty$时，有

$$\begin{cases} f_N(\lambda) \xrightarrow{L^2} f(\lambda), \\ \mathrm{D}(f_N(\lambda)) \to 0. \end{cases} \tag{5.177'}$$

证明 由$|b_N(\lambda)|\to 0$以及(5.175)和(5.54)即可得(5.177)的结论。 ❙

3° 经典谱窗和核函数

为了寻找好的谱窗核函数$K(t)$，许多作者花了很多精力(见[12],[50],[68],[69])。以下介绍一些常见的具代表性的核函数、谱窗和$f_N(\lambda)$。

(1) 矩形窗(或称截尾窗)

$$K(t)=\begin{cases}1, & \text{当}|t|\leqslant 1;\\ 0, & \text{其它。}\end{cases} \tag{5.178'}$$

$$w_N(k)=\begin{cases}1, & |k|\leqslant m_N;\\ 0, & \text{其它。}\end{cases} \tag{5.179'}$$

$$W_N(\lambda)=\frac{1}{2\pi}\left(\frac{\sin\left(m_N+\frac{1}{2}\right)\lambda}{\sin\frac{\lambda}{2}}\right). \tag{5.180'}$$

由(5.6),(5.121)及(5.170)知,

$$f_N(\lambda)=\int_\Pi\left(\frac{1}{2\pi}\sum_{k=1-N}^{N-1}\hat{\gamma}_k e^{-ik\mu}\right)W_N(\lambda-\mu)\,d\mu$$

$$=\frac{1}{2\pi}\sum_{k=1-N}^{N-1}\hat{\gamma}_k\int_\Pi W_N(\lambda-\mu)e^{-ik\mu}d\mu.$$

$$f_N(\lambda)=\frac{1}{2\pi}\sum_{k=1-N}^{N-1}\hat{\gamma}_k w_N(k)e^{-ik\lambda}, \tag{5.181'}$$

其中

$$\hat{\gamma}_k=\frac{1}{N}\sum_{t=1}^{N-k}x_t x_{t+k},\quad k=0,1,\cdots,N-1. \tag{5.182}$$

由(5.179),

$$f_N(\lambda)=\frac{1}{2\pi}\sum_{k=-m_N}^{m_N}\hat{\gamma}_k e^{-ik\lambda}. \tag{5.183}$$

和周期图(5.6)相比，所不同的是求和项数少了(m_N 约$\sqrt{N}$，比$N-1$要少好多)，但(5.183)却是 $f_x(\lambda)$的相容估计。

(2) Bartlett 窗

$$K(t)=\begin{cases}1-|t|, & \text{当}|t|\leqslant 1;\\ 0, & \text{其它。}\end{cases} \tag{5.184}$$

$$w_N(k)=\begin{cases}1-\dfrac{|k|}{m_N}, & \text{当}|k|\leqslant m_N;\\ 0, & \text{其它。}\end{cases} \tag{5.185}$$

$$W_N(\lambda)=\frac{1}{2\pi m_N}\left(\frac{\sin^2\frac{m_N}{2}\lambda}{\sin^2\frac{\lambda}{2}}\right). \tag{5.186}$$

由(5.181)，

$$\hat{f}_N(\lambda)=\frac{1}{2\pi}\sum_{k=-m_N}^{m_N}\left(1-\frac{|k|}{m_N}\right)\hat{\gamma}_k e^{-ik\lambda}. \tag{5.187}$$

(3) Hamming 窗

$$K(t)=\begin{cases}0.54+0.46\cos\pi t, & \text{当}|t|\leqslant 1;\\ 0, & \text{其它。}\end{cases} \tag{5.188}$$

$$w_N(k)=\begin{cases}0.54+0.46\cos\pi\dfrac{k}{m_N}, & \text{当}|k|\leqslant m_N;\\ 0, & \text{其它。}\end{cases} \tag{5.189}$$

$$W_N(\lambda)=\frac{1}{2\pi}\left[0.54\frac{\sin\left(m_N+\frac{1}{2}\right)\lambda}{\sin\frac{1}{2}\lambda}+0.23\left\{\frac{\sin\left(m_N+\frac{1}{2}\right)\left(\lambda+\frac{\pi}{m_N}\right)}{\sin\frac{1}{2}\left(\lambda+\frac{\pi}{m_N}\right)}+\frac{\sin\left(m_N+\frac{1}{2}\right)\left(\lambda-\frac{\pi}{m_N}\right)}{\sin\frac{1}{2}\left(\lambda-\frac{\pi}{m_N}\right)}\right\}\right]. \tag{5.190}$$

$f_N(\lambda)$容易由(5.181)写出。

(4) Hanning 窗

$$K(t)=\begin{cases}\frac{1}{2}(1+\cos\pi t), & |t|\leqslant 1;\\ 0, & 其它。\end{cases}\tag{5.191}$$

$$w_N(k)=\begin{cases}\frac{1}{2}\left(1+\cos\pi\frac{k}{m_N}\right), & |k|\leqslant m_N;\\ 0, & 其它。\end{cases}\tag{5.192}$$

$$W_N(\lambda)=\frac{1}{2\pi}\left[0.5\frac{\sin\left(m_N+\frac{1}{2}\right)\lambda}{\sin\frac{1}{2}\lambda}+0.25\left\{\frac{\sin\left(m_N+\frac{1}{2}\right)\left(\lambda+\frac{\pi}{m_N}\right)}{\sin\frac{1}{2}\left(\lambda+\frac{\pi}{m_N}\right)}+\frac{\sin\left(m_N+\frac{1}{2}\right)\left(\lambda-\frac{\pi}{m_N}\right)}{\sin\frac{1}{2}\left(\lambda-\frac{\pi}{m_N}\right)}\right\}\right].\tag{5.193}$$

(5) Parzen Ⅱ 窗

$$K(t)=\begin{cases}1-6t^2+6|t|^3, & 当|t|\leqslant\frac{1}{2};\\ 2(1-|t|)^3, & \frac{1}{2}\leqslant|t|\leqslant 1;\\ 0, & 其它。\end{cases}\tag{5.194}$$

$$w_N(k)=\begin{cases}1-6\left(\frac{k}{m_N}\right)^2+6\left|\frac{k}{m_N}\right|^3, & |k|\leqslant\frac{1}{2}m_N;\\ 2\left(1-\left|\frac{k}{m_N}\right|\right)^3, & \frac{1}{2}m_N\leqslant|k|\leqslant m_N;\\ 0, & 其它。\end{cases}\tag{5.195}$$

$$W_N(\lambda)\sim\frac{3}{8\pi m_N^3}\left(\frac{\sin(m_N\lambda/4)}{\frac{1}{2}\sin(\lambda/2)}\right)^4.\tag{5.196}$$

(6) Bartlett-Priestley 窗

$$K(t)=\frac{3}{(\pi t)^2}\left(\frac{\sin\pi t}{\pi t}-\cos\pi t\right). \qquad (5.197)$$

$$w_N(k)=\frac{3m_N^2}{(\pi k)^2}\left(\frac{\sin\left(\pi\frac{k}{m_N}\right)}{\pi\frac{k}{m_N}}-\cos\pi\frac{k}{m_N}\right). \qquad (5.198)$$

$$W_N(\lambda)\sim\begin{cases}\frac{3m_N}{4\pi}\left(1-\left(\frac{m_N\lambda}{\pi}\right)^2\right), & |\lambda|\leqslant\pi/m_N;\\ 0, & |\lambda|>\pi/m_N.\end{cases} \qquad (5.199)$$

(7) Daniel 窗

$$K(t)=(\sin\pi t)/(\pi t). \qquad (5.200)$$

$$w_N(k)=\sin\pi\frac{k}{m_N}\Big/\left(\pi\frac{k}{m_N}\right),\ 一切\,k. \qquad (5.201)$$

$$W_N(\lambda)=\begin{cases}\frac{m_N}{2\pi}, & |\lambda|\leqslant\frac{\pi}{m_N};\\ 0, & 其它.\end{cases} \qquad (5.202)$$

必须指出：(6)，(7)两种窗核函数$k(t)$并不满足

$$K(t)=0,\qquad |t|\geqslant 1.$$

事实上，还有许多学者提出的谱窗核函数也不满足这一条件(见[12])。

以上这些窗，哪一个更好？如何判别一个窗比另一个好是一个很复杂的问题，不同的观点有不同的评价方法，有的从偏度$b_N(\lambda)$和方差$D(f_N(\lambda))$来考查(如[68]，[70])有的从L^2误差准则来考查(如[71])或从能量比的准则(如[72])或极大极小准则(如[69])等等。

我们从一个简单的观点来分析，即谱的相对偏差问题。由(5.181)知，当N充分大时，渐近地有

$$D(f_N(\lambda))\sim\frac{m_N}{N}f_\xi^2(\lambda)\|K\|^2,$$

从而相对偏差渐近地为

$$\delta_N(\lambda)=\sqrt{\frac{D(f_N(\lambda))}{f_x^2(\lambda)}}\sim\sqrt{\frac{m_N}{N}}\|K\|. \tag{5.203}$$

如果我们约定取 m_N 为满足(5.167)—(5.168)条件的数列，比如说 $m_N=O(\sqrt{N})$，则(5.203)表明相对偏差 $\delta_N(\lambda)$ 与 $\|K\|$ 是成正比的，因此无疑，好的谱窗核函数应使得 $\|K\|$ 尽可能地小(当然这不是唯一指标)。表5.1给出了前述七种核函数的核能 $\|K\|^2$。

表 5.1

谱窗名称	窗核函数 $K(t)$	核能 $\|K\|^2$
矩形	$K(t)=\begin{cases}1, & \|t\|\leqslant 1\\ 0, & \|t\|>1\end{cases}$	2
Bartlett	$K(t)=\begin{cases}1-\|t\|, & \|t\|\leqslant 1\\ 0, & \text{其它}\end{cases}$	$\frac{2}{3}=0.6667$
Daniel	$K(t)=(\sin\pi t/\pi t)$	1
Hamming	$K(t)=\begin{cases}0.54+0.46\cos\pi t, & \|t\|\leqslant 1\\ 0, & \|t\|>1\end{cases}$	0.7948
Hanning	$K(t)=\begin{cases}\frac{1}{2}(1+\cos\pi t), & \|t\|\leqslant 1\\ 0, & \|t\|>1\end{cases}$	$\frac{3}{4}=0.75$
Parzen Ⅱ	$K(t)=\begin{cases}1-6t^2+6\|t\|^3, & \|t\|\leqslant 1/2\\ 2(1-\|t\|)^3, & \frac{1}{2}\leqslant\|t\|\leqslant 1\\ 0, & \|t\|>1\end{cases}$	0.53928
Priestley	$K(t)=\frac{3}{\pi t^2}\left(\frac{\sin\pi t}{\pi t}-\cos\pi t\right)$	$\frac{6}{5}=1.2$

以上各种窗中，以ParzenⅡ的核能最小，$\|K\|^2=0.53928$，而以矩形窗最大 $\|K\|^2=2$。由(5.203)观察即：在相同的样本数 N 和滞后值 m_N 之下，用矩形窗比用 ParzenⅡ 窗的相对估计方差要大

3.7倍。

但Parzen Ⅱ 窗的缺点是核函数表达式比较繁，所以在工程上不如 Hanning 窗那样被广泛使用。[72]介绍了从能量比观点下导出的近似最优谱窗核函数

$$K(t)=\begin{cases}0.313+0.531\cos\pi t+0.156\cos 2\pi t, & |t|\leqslant 1;\\ 0, & |t|>1.\end{cases}\tag{5.204}$$

此时，不仅谱窗函数 $W_N(\lambda)$ 能量集中，而且核能是目前最小的，

$$\|K\|^2=0.502.\tag{5.205}$$

4° 加窗谱估计的实际计算

数学上不难证明有如下的有限 Fourier 变换的结果：

设 f_k 为$[0,2\pi]$上等间距$\frac{2k\pi}{M}$ 点 $f(x)$的函数值，记

$$c_l=\frac{1}{M}\sum_{k=0}^{N-1}f\left(\frac{2k\pi}{M}\right)e^{-ikl\frac{2\pi}{M}},\tag{5.206}$$

$$l=0,1,2,\cdots,M-1$$

则

$$f_k=\sum_{l=0}^{M-1}c_l\,e^{ikl\frac{2\pi}{M}},\qquad k=0,1,2,\cdots,M-1.\tag{5.207}$$

表明了$\{c_l\}$和$\{f_k\}$是一对有限离散的 Fourier 变换。

再考察(5.181)的 $f_N(\lambda)$。由于通常选择的核$K(t)$具有性质

$$K(t)=0,\qquad |t|>1,$$

从而

$$w_N(k)=0,\qquad |t|>m_N.$$

所以(5.181)变成

$$\begin{aligned}f_N(\lambda)&=\frac{1}{2\pi}\sum_{l=-m_N}^{m_N}\gamma_l w_N(l)e^{-il\lambda}\\&=\frac{1}{\pi}\left\{\frac{1}{2}\gamma_0+\sum_{l=1}^{m_N}\gamma_l w_N(l)\cos l\lambda\right\}.\end{aligned}\tag{5.208}$$

表明$f_N(\lambda)$是由m_N+1个值$\{\hat{\gamma}_l w_N(l)\}_0^{m_N}$所确定，因此从有限Fourier变换的对偶性观点看，$\{\hat{\gamma}_l w_N(l)\}_0^{m_N}$亦可由$f_N(\lambda_k)$的$m_N+1$个值$\left\{f_N\left(\frac{k\pi}{m_N}\right)\right\}_0^{m_N}$来确定①。因此，在时间序列的实际计算中多采用以下的计算公式

$$\begin{cases} \hat{\gamma}_k = \dfrac{1}{N}\sum\limits_{l=1}^{N-k} x_l\, x_{l+k}; \\ \hat{f}_k = f_N\left(\dfrac{k\pi}{m_N}\right) = \dfrac{1}{\pi}\left\{\dfrac{\hat{\gamma}_0}{2} + \sum\limits_{l=1}^{m_N} \hat{\gamma}_l w_N(l)\cos\left(l\dfrac{k\pi}{m_N}\right)\right\}, \end{cases} \tag{5.209}$$

$$k=0,1,2,\cdots,m_N.$$

从保存信息观点看，(5.208)只为m_N+1个$\{\hat{f}_k\}_0^{m_N}$所确定。所以通常并不计算$f_N(\lambda),\lambda\in\Pi$的全部函数值而只算$\{\hat{f}_k\}_0^{m_N}$。当然也没有理由说不能用(5.208)算其它$\lambda\in\Pi$的谱值(见[50])。事实上以往的谱估计的偏度和方差的估计都是对$\lambda\in[-\pi,\pi]$而言的。

根据(5.209)一种实际估计谱密度的步骤可按以下进行：

(1) 先估算在$[0,\pi]$上希望计算谱线$\{\hat{f}_k\}$的总数m_N+1，它将决定对$f_x(\lambda)$的分辨度。过小的m_N将遗漏对$f_x(\lambda)$变化的认识，过多则不仅浪费计算时间，而且主要的，以下将看到将加大对x_t的观测样本N。当然实用中也有用实验方法来确定m_N的。

(2) 在m_N确定之后，选择一种谱窗核$K(t)$，由表5.1可确定相应的核能$\|K\|^2$。再由工作的需要，确定一个相对误差δ_N(见(5.203))，于是可导出

$$N\sim\frac{m_N\|K\|^2}{\delta_N^2}, \tag{5.210}$$

由此确定应观测的样本值N。通常δ_N为0.10～0.30左右。误差愈小，样本N愈大。要求对$f_x(\lambda)$分辨愈细，m_N愈大，则样本也愈多——过大的N，即过长的时间观测有时就不能保持x_t的

① 因为$f_N(-\lambda)=f_N(\lambda)$。

平稳性，得到的 $\hat{f}_N(\lambda)$ 也就不可靠了。

例如：要求 $m_N=30$，$\delta_N=0.30$，选择 Bartlett 窗。这时 $\|K\|_B^2=\frac{2}{3}$，因此确定的样本数大致为

$$N\sim\frac{30\times 0.6667}{0.09}=222.$$

若选 Parzen Ⅱ，则 $N\sim 180$；矩形窗 $N\sim 667$。

（3）对观测资料 $x_1,x_2,\cdots,x_N$，先中心化

$$y_l=x_l-\bar{x},\qquad l=1,2,\cdots,N,$$

然后计算

$$\hat{\gamma}_k=\frac{1}{N}\sum_{l=1}^{N-k}y_l\,y_{l+k},\quad k=0,1,\cdots,m_N.$$

（4）计算

$$\hat{f}_k=\frac{1}{\pi}\left(\frac{\hat{\gamma}_0}{2}+\sum_{l=1}^{m_N}\hat{\gamma}_l w_N(l)\cos\left(\frac{lk\pi}{m_N}\right)\right),\qquad k=0,1,\cdots,m_N.$$

最后我们指出：如果 x_t 序列是由连续过程抽样得来，则按第一章中介绍的有关知识（见定理1.15），需考虑过程的一个上界频，设为 W，则采样间隔为 $\Delta t=\frac{1}{2W}$，即

$$x_k=x(k\cdot\Delta t),\qquad k=0,1,2,\cdots,N,\tag{5.211}$$

然后进行以上的（1）—（4）的计算步骤，这时算出的 $\{\hat{f}_k,k=0,1,\cdots,m_N\}$ 即相当于分布在 $[0,2\pi W]$ 上的原连续过程（等距）的谱密度，差一个倍数 Δt 因子，即

$$f_x(\Omega_k)=\Delta t\hat{f}_k,\ \Omega_k\in[0,2\pi W],\ k=0,1,\cdots,m_N.\tag{5.212}$$

§4 离散周期谱的检测

§3中介绍的是当 x_t 具有谱密度时如何从观测样本 $x_1,x_2,\cdots,x_N$ 去进行谱密度估计。在应用中一类非常重要的问题是在具

有背景噪声的条件下如何发现离散的周期谱，这时可分为两种类型：

$$(1)\ x_t=\sum_{k=1}^{p} A_k e^{it\lambda_k}+\xi_t, \tag{5.213}$$

其中 ξ_t 是平稳序列，A_k,λ_k 皆为常数，所以这种问题是具有确定型分量的离散谱的检测问题。

$$(2)\ x_t=\sum_{k=1}^{p} \eta_k e^{i\lambda_k t}+\xi_t=\eta(t)+\xi_t, \tag{5.214}$$

其中 $\{\eta_k\}_1^p$ 是随机变量，ξ_t 为平稳序列。如果能对离散谱 $F_\eta(\lambda)$ 和谱密度 $f_\xi(\lambda)$ 进行估计，则 $F_x(\lambda)$ 就是混合谱，因而也可以得到估计。

以下先讨论第一类问题，这类问题实际上很早就有人讨论了如[68],[73],[74]。

1° 白噪声背景下对确定型离散谱的统计检测

以下介绍的是应用统计学中的假设检验的方法来检测 x_t（见(5.213)式）中的 (A_k,λ_k)，此时假定 ξ_t 是正态白噪声。

设 $x_1,x_2,\cdots,x_N$ 是正态白噪声的观测，令

$$A(\lambda)=\sqrt{\frac{2}{N}}\sum_{k=1}^{N} x_k\cos k\lambda, \tag{5.215}$$

$$B(\lambda)=\sqrt{\frac{2}{N}}\sum_{k=1}^{N} x_k\sin k\lambda. \tag{5.216}$$

则

$$\begin{aligned} I_N(\lambda)&=A^2(\lambda)+B^2(\lambda)\\ &=\frac{2}{N}\left\{\left(\sum_{k=1}^{N}x_k\cos k\lambda\right)^2+\left(\sum_{k=1}^{N}x_k\sin k\lambda\right)^2\right\}\\ &=4\pi I_N(\lambda), \end{aligned} \tag{5.217}$$

其中 $I_N(\lambda)$ 是以往引入的周期图。

以下不妨假定 $N=2M$ 是正偶数，并假定

$$x_t \text{ 为i.i.d.} \sim N(0,\sigma^2). \tag{5.218}$$

于是

$$D(A(\lambda)) = D\left(\sqrt{\frac{2}{N}}\sum_{k=1}^{N} x_k \cos k\lambda\right)$$

$$= \frac{2}{N}\sigma^2\left(\sum_{k=1}^{N}\cos^2 k\lambda\right). \tag{5.219}$$

但

$$\sum_{k=1}^{N}\cos^2 k\lambda = \frac{N}{2} + e^{2i\lambda}\frac{1-e^{2iN\lambda}}{1-e^{2i\lambda}} + e^{-2i\lambda}\frac{1-e^{-2iN\lambda}}{1-e^{-2i\lambda}} \tag{5.220}$$

$$= \begin{cases} \dfrac{N}{2}, & \text{当 } \lambda = \dfrac{2p\pi}{N},\ p \neq 0, \dfrac{N}{2};\ 0 < p < \dfrac{N}{2}; \\ N, & \text{当 } \lambda = 0, \pm\pi\ (p \text{ 为整数}), \end{cases} \tag{5.221}$$

故

$$D(A(\lambda)) = \begin{cases} \sigma^2, & \lambda = p\dfrac{2\pi}{N},\ 0 < p < \dfrac{N}{2},\ p \neq 0,\ \dfrac{N}{2}; \\ 2\sigma^2, & \lambda = p\dfrac{2\pi}{N},\ p = 0 \text{ 或 } \dfrac{N}{2}. \end{cases} \tag{5.222}$$

可见

$$\begin{cases} A(\lambda_p) \sim N(0,\sigma^2), & \text{当 } p \neq 0, \dfrac{N}{2},\ 0 < p < \dfrac{N}{2}; \\ A(\lambda_p) \sim N(0,2\sigma^2), & \text{当 } p = 0, \dfrac{N}{2}. \end{cases} \tag{5.223}$$

不难看出

$$B(\lambda_p) \sim \begin{cases} N(0,\sigma^2), & \text{当 } p \neq 0,\ \dfrac{N}{2}; \\ N(0,2\sigma^2), & \text{当 } p = 0,\ \dfrac{N}{2}. \end{cases} \tag{5.224}$$

$$\begin{aligned} &\mathrm{Cov}(A(\lambda_p),\ B(\lambda_q)) \\ &= \sigma^2\frac{2}{N}\sum_{k=1}^{N}\cos(\lambda_p k)\sin(\lambda_q \cdot k) \\ &= 0, \quad \text{对一切 } p, q, \end{aligned} \tag{5.225}$$

$$\mathrm{Cov}(A(\lambda_p),A(\lambda q))=\mathrm{Cov}(B(\lambda_p),B(\lambda q))$$
$$=0, \quad 当\ p\neq q。\tag{5.226}$$

因此当 $p\neq 0,\dfrac{N}{2}$时，$\hat{I}_N(\lambda_p)=A^2(\lambda p)+B^2(\lambda p)$遵从 χ^2 分布律，即有：

定理5.9 设 x_t 是一串相互独立同 $N(0,\sigma^2)$分布的序列，$A(\lambda_p),B(\lambda_q)$由(5.215)和(5.216)定义，$\lambda_p=\dfrac{2p\pi}{N}$。则

(1) 当 $p\neq 0,\dfrac{N}{2}$，$\hat{I}_N(\lambda_p)\sim\sigma^2\chi_2^2$ 分布；　　(5.227)

(2) 当 $p=0,\dfrac{N}{2}$，$\hat{I}_N(\lambda_p)\sim(2\sigma^2)\chi_1^2$ 分布；　　(5.228)

(3) 对一切 p,q $A(\lambda_p)$，$B(\lambda_q)$相互独立；

(4) $\{A(\lambda_p)\}$和$\{B(\lambda p)\}$ i.i.d. $\sim N(0,\sigma_p^2)$分布，其中 $\sigma_p^2=\sigma^2$，若 $p\neq 0,\dfrac{N}{2}$；$\sigma_p^2=2\sigma^2$，若 $p=0,\dfrac{N}{2}$。

由于χ_r^2变量的期望是 r，方差为 $2r$，故有

推论 在定理5.9条件下

$$E[\hat{I}(\lambda_p)]=2\sigma^2, \quad p=0,1,\cdots,\frac{N}{2}。\tag{5.229}$$

$$D(\hat{I}(\lambda_p))=\begin{cases}4\sigma^4, & p\neq 0,\dfrac{N}{2};\\ 8\sigma^4, & p=0,\dfrac{N}{2}。\end{cases}\tag{5.230}$$

证明

$$E[I(\lambda_p)]=\begin{cases}\sigma^2E\chi_2^2=2\sigma^2, & p\neq 0,\dfrac{N}{2};\\ 2\sigma^2E\chi_1^2=2\sigma^2, & p=0,\dfrac{N}{2}。\end{cases}$$

$$D(I(\lambda_p)) = \begin{cases} \sigma^4 D(\chi_2^2) = 4\sigma^4 , & p \neq 0, \dfrac{N}{2}; \\ (2\sigma^2)^2 D(\chi_1^2) = 8\sigma^4, & p = 0, \dfrac{N}{2}. \end{cases}$$ ▌

以下来考虑(5.213)式的模型下的检测问题，

即

$$x_t = \sum_{k=1}^{P} A_k \cos(\omega_k t) + \xi_t, \tag{5.231}$$

其中 P 已知，A_k,ω_k 为常数(未知)，$\xi_t \sim N(0,\sigma^2)$ i.i.d.，但 σ^2 未知。考虑统计量

$$\hat{I}_N(\lambda) = \frac{2}{N} \left| \sum_{k=1}^{N} x_k e^{-ik\lambda} \right|^2.$$

为了检测(ω_k)和(A_k)，可以作以下的假设检验：

$$H_0:\ A_i \equiv 0, \quad i = 1,2,\cdots,P. \tag{5.232}$$

选统计量为 $\hat{I}_N(\lambda)$。于是 H_0 成立意味着 $x_t \equiv \xi_t$。简记

$$\hat{I}_P = \hat{I}_N\left(\frac{2p\pi}{N}\right). \tag{5.233}$$

为方便起见，取 $N = 2M+1$ 为奇数，$\left[\dfrac{N}{2}\right] = M$。于是当 $p \neq 0$，$\hat{I}_p \sim \sigma^2 \cdot \chi_2^2$ 分布，或

$$\tilde{I}_P = \frac{\hat{I}_P}{\sigma^2} \sim \chi_2^2, \tag{5.234}$$

即

$$P\{\tilde{I}_P \leqslant z\} = \int_0^z p_{\tilde{I}_p}(x)\,dx = 1 - e^{-\frac{z}{2}}. \tag{5.235}$$

记

$$\gamma = \frac{\max\limits_{1\leqslant p\leqslant M} \{I_p\}}{\sigma_x^2}. \tag{5.236}$$

在 H_0 之下，$\sigma_x^2 = \sigma^2$。显见

$$P\{\gamma > z\} = 1 - P\{\gamma \leqslant z\}$$

$$= 1 - P\{\tilde{I}_p \leqslant z,\ \text{一切}\ p\}$$
$$= 1 - (1 - e^{-\frac{z}{2}})^M, \qquad (5.237)$$

其中最后一个等式是利用了$\{\tilde{I}_p\}$的独立性。

由上可见，如果σ_z^2已知，则利用γ作统计量和分布(5.237)即可对H_0作假设检验。但问题恰好是在实际工作中σ_z^2是未知的，因而需要加以估计。因此把问题作以下更动：上一段已证明了，若x_t是纯噪声，即$N(0,\sigma^2)$i.i.d.变量，则

$$E[\hat{I}(\lambda_p)] = 2\sigma_z^2 \quad \left(0 \leqslant p \leqslant \frac{N}{2}\right)$$

(见(5.229))，于是把H_0换成

$$H_0:\ E[\hat{I}_p] = 2\sigma_z^2\ (\text{一切}\ p \geqslant 1). \qquad (5.238)$$

由于

$$E\left[\sum_{p=1}^{[N/2]} \hat{I}_p\right] = 2\sigma_z^2 \left[\frac{N}{2}\right] \qquad (5.239)$$

所以

$$v = \frac{\sum_{p=1}^{[N/2]} \hat{I}_p}{2\left[\frac{N}{2}\right]} \qquad (5.240)$$

的偏度为

$$Ev = \sigma_z^2, \qquad (5.241)$$

因而v就可用来作为σ_z^2的估计。由此将(5.236)的γ作如下修改(相当于t检验的“学生化”)

$$g^* = \frac{\max(\hat{I}_p)}{\left(\frac{1}{2[N/2]}\right)\sum_{p=1}^{[N/2]} \hat{I}_p}. \qquad (5.242)$$

Fisher (1929)年得到确切的

$$g = \frac{g^*}{2[N/2]} = \frac{\max(\hat{I}_p)}{\sum_{p=1}^{[N/2]} \hat{I}_p} \qquad (5.243)$$

的分布：

$$P\{g>z\}=M(1-z)^{M-1}-\frac{M(M-1)}{2}(1-2z)^{M-1}+\cdots$$

$$+(-1)^{a}\frac{M!}{a!(M-a)!}(1-az)^{M-1}, \qquad (5.244)$$

其中

$$M=\left[\frac{N}{2}\right], \qquad a=\left[\frac{1}{z}\right].$$

利用 $g^*=2Mg$ 容易得到 Fisher 分布的推论：

$$P\{g^*>z\}=P\left\{g>\frac{z}{2M}\right\}\sim M\left(1-\frac{z}{2M}\right)^{M-1} \qquad (5.245)$$

$$\sim Me^{-z/2}. \qquad (5.246)$$

其实由 g^* 的(5.242)近似看 ν 为 σ_x^2 的估计，由 γ 的概率分布(5.237)，则

$$P\{g^*>z\}\sim P\{\gamma>z\}=1-(1-e^{-\frac{z}{2}})^{M}\sim Me^{-\frac{z}{2}} \qquad (5.247)$$

与(5.246)类似。

Grenander 与 Rosenblatt(1957)(见[68])提出如下的假设检验：

$$H_0:\ x_t=\xi_t,\ \text{i.i.d.}\ N(0,\sigma^2).$$

统计量的选择：记$\{\hat{I}_p\}$中第 r 个最大的为 $\hat{I}_{(r)}$，令

$$g_{(r)}=\frac{\hat{I}_{(r)}}{\sum_{p=1}^{M}\hat{I}_p}, \qquad (5.248)$$

则 Grenander 和 Rosenblatt 证明了

$$P\{g_{(r)}>z\}=\frac{M!}{(r-1)!}\sum_{j=1}^{a}\frac{(-1)^{j-r}(1-jz)^{M-1}}{j(M-j)!(j-r)!}, \quad (5.249)$$

其中

$$M=[N/2],\ a=[1/z],$$

Shimshoni, M.(1971)于[75]中对 $r=1,2,5,7,10,25,50$等给

出了在 α 水平下检验的分位点。

假若通过检验而推翻 H_0，则表明 x_t 不是纯白噪声而是含有 r 个频率分量（此处 r 是先验指定）。$\lambda_p = p\frac{2\pi}{N}$，$p$ 则对应于 $\hat{I}_{(r)}$ 大小的头 r 个值，记为 $\lambda_{p_1}^{(N)}, \lambda_{p_2}^{(N)}, \cdots, \lambda_{p_r}^{(N)}$，于是对应于(5.213)的确定型分量的振幅 $\{A_k\}_1^r$ 可用

$$\hat{A}_k^{(N)} = \frac{1}{N}\sum_{n=1}^{N} x_n e^{-in\lambda_{p_k}^{(N)}}, \quad k = 1, 2, \cdots, r \quad (5.250)$$

来估计。

最后提醒注意的是：我们不能简单地对 $g_{(1)}$，$g_{(2)}$，$\cdots$，$g_{(r)}$ 逐个作检验来确定 r 值，因为如果 $r=1$（最大）检验后推翻了 H_0，则表明 $x_t \neq \xi_t$（白噪声），其中已有周期分量，在这样的条件下 $g_{(2)}$ 的分布已起了变化。

2° 混合谱的估计，HYS 方法

以上讨论了白噪声背景下运用假设检验方法对离散谱的检测问题，以下我们来讨论更一般的(5.99)的线性模型下对离散谱的估计。

设

$$x_t = \sum_{k=1}^{P} \eta_k e^{it\lambda_k} + \xi_t, \quad (5.99)$$

其中 ξ_t 满足弱 P 条件，η_k 是非零有界(a.s.)的变量，λ_k 是未知的常数，P 亦未知。

首先介绍[76]中对阶数 P 的估计方法。

定义5.2 设 $\{g_N(\lambda)\}$ 是 Π 上的一列实值连续函数，γ 是给定的正数，$D_N \subset \Pi$ 是一个区间，

$$D_N = \begin{cases} (\alpha, \beta), & (\alpha, \beta) \subset \Pi; \\ [-\pi, \alpha) \cup (\beta, \pi]. & \end{cases} \quad (5.251)$$

如果它满足以下三条件，就称为是 $g_N(\lambda)$ 的一个 γ 区间：

(1) 对 $\lambda\in D_N$， $g_N(\lambda)<\gamma$； (5.252)

(2) $g_N(\alpha)=g_N(\beta)=\gamma$； (5.253)

(3) $\mu(D_N)>\frac{1}{\sqrt{N}}$ (Lebesgue 测度)。 (5.254)

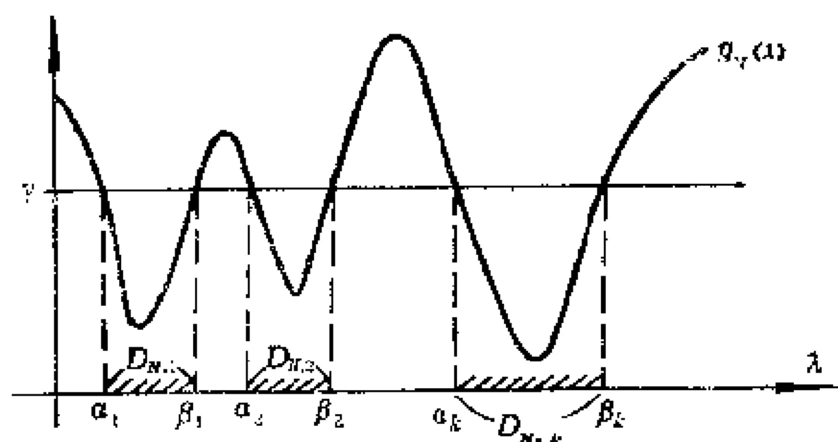


图 5.4

图5.4中$(\alpha_i,\beta_i)=D_{N,i}$皆为$\gamma$区间。

定义5.3 设x_t满足模型(5.99)，ξ_t是概率空间$(\Omega,\mathscr{F},P)$上的弱P平稳列，对$\gamma>0$，$\omega\in\Omega$，$x_1(\omega),x_2(\omega),\cdots,x_N(\omega)$为$x_t$的样本，将$I_{xN}(\lambda)$(见(5.102)式)的$\gamma$区间的个数记为$\hat{P}_N(\omega)$(简记为$\hat{P}_N$)。

显见，对$\gamma>0$，由γ区间的定义，各γ区间必互不相交并只有有限个。于是当N充分大时以下定理说明$\hat{P}_N$是(5.99)模型真阶的强相容估计。

定理5.10 设x_t为(5.99)的线性模型，则

$$\lim_{N\to\infty}\hat{P}_N(\omega)=P\quad \text{a.s.}, \tag{5.255}$$

其中P为模型中离散部分的真阶。

证明 对诸频率$\{\lambda_j\}_1^p$任选一个j，记

$$H_j=[\lambda_j+N^{-\frac{15}{16}},\ \lambda_{j+1}-N^{-\frac{15}{16}}]. \tag{5.256}$$

对$\lambda\in H_j$，则由$\mathscr{A}$(见(5.105))之定义知$\lambda\in\mathscr{A}$，于是由定

理5.7的(5.104)知，Pr.1 地当 N 充分大之后有

$$\sup_{\lambda\in H_j} J_{xN}(\lambda)\leqslant CN^{-\frac{1}{16}}<\gamma. \tag{5.257}$$

这时

$$\mu(H_j)=(\lambda_{j+1}-\lambda_j)-2N^{-\frac{15}{16}}. \tag{5.258}$$

而由(5.109)知

$$(\lambda_{j+1}-\lambda_j)\geqslant 4\delta_0, \tag{5.259}$$

故

$$\mu(H_j)\geqslant 4\delta_0-2N^{-\frac{15}{16}}>\frac{1}{\sqrt{N}}. \tag{5.260}$$

另一方面，对 $j=1,2,\cdots,P$，由定理5.7的(5.103)，

$$J_{xN}(\lambda_j)\geqslant c_0N^{\frac{1}{16}}>\gamma. \tag{5.261}$$

由连续函数的性质，必有 $J_{xN}(\lambda)$ 的一个 γ 区间包含在 $[\lambda_j,\lambda_{j+1}]$ 之间，记为

$$D_{j,N}=(b_{j,N},\ a_{j+1,N}),\qquad j=1,2,\cdots,P, \tag{5.262}$$

$$b_{j,N}-a_{j,N}<2N^{-\frac{15}{16}}<\frac{1}{\sqrt{N}}. \tag{5.263}$$

而

$$a_{j+1,N}-b_{j,N}>\min_j\ \mu(H_j)>\frac{1}{\sqrt{N}}$$

(对端点 $[\lambda_P,\pi]\cup[-\pi,\lambda_1]$ 可作类似讨论)。因此可得

$$a_{1,N}<\lambda_1<b_{1,N}<a_{2,N}<\lambda_2<\cdots<a_{P,N}<\lambda_P<b_{P,N}<a_{1,N}+2\pi.$$

可见 P_r 1 地当 N 充分大后 $J_{xN}(\lambda)$ 恰好有 P 个 γ 区间，即

$$\lim_{N\to\infty}\hat{p}_N(\omega)=P\quad \text{a.s.}. \quad \blacksquare$$

由以上证明过程可得：

推论 1 设 x_t 满足模型(5.99)，对给定的 $\gamma>0$，概率为 1 地当 N 充分大以后，$J_{xN}(\lambda)$ 有 P 个 γ 区间

$$\begin{cases} D_{j,N}=(b_{j,N},a_{j+1,N}), & j=1,2,\cdots,P-1; \\ D_{P,N}=(b_{P,N},\pi]\cup[-\pi,a_{1,N}), \end{cases} \tag{5.264}$$

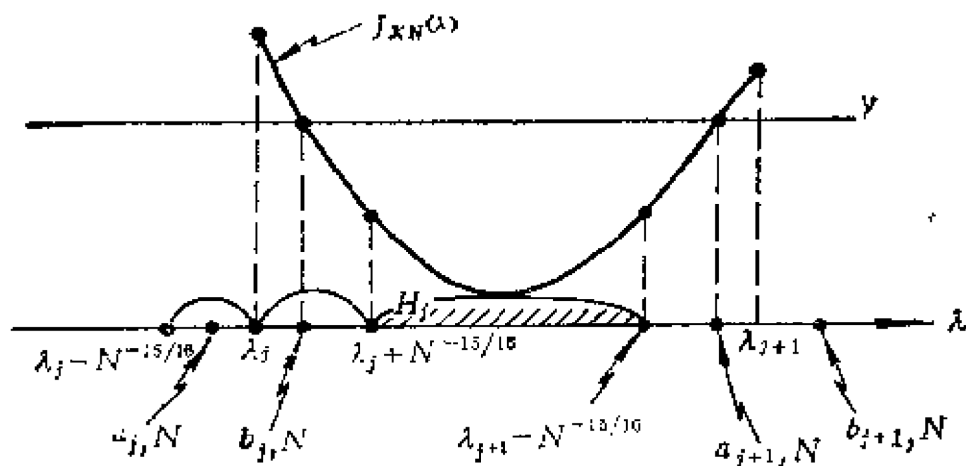


图 5.5

其中$(a_{j,N}, b_{j,N})$满足(5.263)，并且

$$0 \leqslant \lambda_j - a_{j,N} < N^{-15/16}, \quad j=1,2,\cdots,P, \tag{5.265}$$

$$0 \leqslant b_{j,N} - \lambda_j < N^{-15/16}, \tag{5.266}$$

$$a_{j+1,N} - b_{j,N} > 2\delta_0, \tag{5.267}$$

$$2\pi - (b_{P,N} - a_{1,N}) > 2\delta_0. \tag{5.268}$$

推论 2 设 x_t 满足模型(5.99)。对任何

$$\beta < \frac{15}{16}, \quad \hat{\lambda}_{j,N} \in [a_{j,N}, b_{j,N}] 有$$

$$\lim_{N\to\infty} N^{\beta}(\hat{\lambda}_{j,N} - \lambda_j) = 0 \quad \text{a.s.}. \tag{5.269}$$

证明 由定理5.10证明中的(5.263)知

$$b_{j,N} - a_{j,N} < 2N^{-15/16}.$$

故

$$|\hat{\lambda}_{j,N} - \lambda_j| < b_{j,N} - a_{j,N} < 2N^{-15/16}.$$

当 $\beta < \frac{15}{16}$时，设 $\beta = \frac{15}{16} - \varepsilon$，

$$N^{\beta}|\hat{\lambda}_{j,N} - \lambda_j| < 2N^{\frac{15}{16}-\varepsilon}N^{-\frac{15}{16}} = 2N^{-\varepsilon} \to 0. \quad \blacksquare \tag{5.270}$$

为了得到更精确的对$\{\lambda_j\}_1^P$的估计，[76]指出：设 x_t 的周期图 $I_N(\lambda)$在

$$D_{j,N}^{*}=[a_{j,N},\ b_{j,N}],\qquad j=1,2,\cdots,\hat{P}_N \quad (5.271)$$

上的最大值点是 $\lambda_{j,N}^{*}$，则有：

定理5.11 在定理5.10的条件下，对 $\beta<\frac{17}{16}$，有

$$\lim_{N\to\infty} N^{\beta}(\lambda_{j,N}^{*}-\lambda_j)=0 \quad \text{a.s.}. \qquad (5.272)$$

定理5.12 设 x_t 满足定理5.10条件，$\hat{P}_N$ 是(5.255)确定的阶数 P 的强相容估计，$\{\lambda_{j,N}^{*}\}$ 是(5.272)式中的估计，令

$$\hat{a}_{j,N}=\frac{1}{N}\sum_{n=1}^{N}x_n e^{-in\lambda_{j,N}^{*}},\qquad j=1,2,\cdots,\hat{P}_N,\quad (5.273)$$

则对 $\beta<\frac{1}{16}$，有

$$\lim_{N\to\infty} N^{\beta}(\hat{a}_{j,N}-\eta_j)=0 \quad \text{a.s.}, \qquad (5.274)$$

其中 $\{\eta_j\}$ 是模型(5.99)中的随机变量。

需要指出的是：模型(5.99)实质上包含了应用中最有代表性的如下模型

$$x_t=\sum_{j=1}^{P}A_j\cos(\omega_j t+\varphi_j)+\varepsilon_t, \qquad (5.275)$$

其中 P, A_j, ω_j 均为未知的常数，φ_j 是 i.i.d. $U[-\pi,\pi]$ 的随机变量，ε_t 是弱 P 平稳列，例如白噪声或平稳 Gauss 噪声。

事实上可将(5.275)改写为(5.99)形式：

$$x_t=\sum_{j=1}^{2P}\eta_j e^{it\lambda_j}+\xi_t, \qquad (5.276)$$

其中

$$\lambda_j=\begin{cases}\omega_j & j=1,2,\cdots,P;\\ -\omega_{j-P}, & j=P+1,\cdots,2P,\end{cases} \qquad (5.277)$$

$$\eta_j=\begin{cases}\dfrac{1}{2}A_j e^{i\varphi_j}, & j=1,2,\cdots,P;\\[2ex] \dfrac{1}{2}A_{j-P}e^{-i\varphi_{j-P}}, & j=P+1,\cdots,2P,\end{cases} \qquad (5.278)$$

并以 $2|\eta_j|$ 来估 A_j，以 $\arg(\eta_j)$ 来估 φ_j—— 能估随机相位是这一方法的一大优点。

最后指出：在模型

$$x_t = \sum_{k=1}^{p} \eta_k e^{it\lambda_k} + \xi_t$$

中，一旦估出 $\hat{p}_N, \{\hat{\alpha}_{j,N}\}$ 和 $\{\lambda^*_{j,N}\}$ 之后，可令

$$y_t = x_t - \sum_{k=1}^{\hat{p}_N} \hat{\alpha}_{j,N} e^{it\lambda^*_{k,N}}, \tag{5.279}$$

并求 y_t 的一组样本 $\{y_t\}$——于是可以运用第四章和本章介绍过的估计谱密度的方法求 y_t 的谱估计，设为 $\hat{f}_N(\lambda)$，则

$$\hat{F}_N(\lambda) = \sum_{\{k: \lambda^*_{k,N} < \lambda\}} \hat{\alpha}_{k,N} + \int_{-\pi}^{\lambda} f_N(u)\,du, \quad \lambda \in \Pi \tag{5.280}$$

即为 x_t 的混合谱函数的估计。

例5.1 设

$$x_n = \sum_{j=1}^{7} \alpha_j e^{-in\lambda_j} + \xi_n, \quad n = 1, 2, \cdots, N,$$

其中 $\xi_n \sim N(0, 101)$，并为如下的 MA 序列

$$\xi_n = 8\varepsilon_n - 6\varepsilon_{n-1} + \varepsilon_{n-2}, \tag{5.281}$$

$\{\varepsilon_n\}$ 为白噪声。样本 $N = 50, 100$ 时用前述方法通过 Monte-Carlo 模拟可信得如下结果：（选 $\gamma = 7.2$）$\mathrm{SNR} = \sum_{j=1}^{7} \alpha_j^2/\sigma^2 = 1.5$。

N	$\hat{P}_N$	$\lambda^*_{1,N}$	$\lambda^*_{2,N}$	$\lambda^*_{3,N}$	$\lambda^*_{4,N}$	$\lambda^*_{5,N}$	$\lambda^*_{6,N}$	$\lambda^*_{7,N}$
50	6	-2.7646	-2.0734	-1.7592	-1.0053	1.6964	1.6964	—
100	7	-2.7646	-2.1048	-1.7592	-1.0053	1.7278	1.7278	2.0106
真值	7	-2.76	-2.12	-1.76	-1	1.724	1.724	2

$\hat{a}_{1,N}$	$\hat{a}_{2,N}$	$\hat{a}_{3,N}$	$\hat{a}_{4,N}$	$\hat{a}_{5,N}$	$\hat{a}_{6,N}$	$\hat{a}_{7,N}$
5.733	0.775	5.009	−3.618	2.649	—	−3.723
5.403	3.795	−3.859	−4.02	3.126	1.87	−3.283
5.1	6.7	−4.8	−3.7	3.9	3	−4.23

例5.2 $x_n = \sum_{j=1}^{6} a_j e^{-in\lambda_j} + \xi_n$,其中 ξ_n 是 ARMA(3,3)模型

$$3\xi_n + 3.5\xi_{n-1} + \xi_{n-2} = 32\varepsilon_n - 24\varepsilon_{n-1} + 4\varepsilon_{n-2}.$$

$N = 50$ 和 100, 取 $\gamma = 9.6$.

$$\text{SNR} = \sum_{j=1}^{p} a_j^2/\sigma^2 = \frac{102.8}{222} = 0.463.$$

N	$\hat{P}_N$	$\lambda^*_{1,N}$	$\lambda^*_{2,N}$	$\lambda^*_{3,N}$	$\lambda^*_{4,N}$	$\lambda^*_{5,N}$	$\lambda^*_{6,N}$
50	6	−3.1415	−2.0743	−0.7539	0.3141	1.2566	2.3247
100	6	−3.1101	−2.0734	−0.8168	0.3141	1.286	2.2933
真　值	6	−3.1	−2.07	−0.8	0.3	1.29	2.3

$\hat{a}_{1,N}$	$\hat{a}_{2,N}$	$\hat{a}_{3,N}$	$\hat{a}_{4,N}$	$\hat{a}_{5,N}$	$\hat{a}_{6,N}$
−1.64	−3.669	4.557	4.395	−2.938	2.120
−4	−4.2	3.862	3.377	−4.869	3.272
−4.8	−3.9	4.2	3.7	−4.23	3.92

习　　题

1. 设 x_t 是 i.i.d. $N(0,\sigma^2)$ 序列，利用第二章习题 7 中的 Bartlett 公式直接证明

$$\operatorname{Cov}\{\hat{I}_N(\lambda_1),\hat{I}_N(\lambda_2)\}=\frac{4\sigma^4}{N^2}\left\{\frac{\sin^2\frac{N(\lambda_1+\lambda_2)}{2}}{\sin^2\frac{\lambda_1+\lambda_2}{2}}+\frac{\sin^2\frac{N(\lambda_1-\lambda_2)}{2}}{\sin^2\frac{\lambda_1-\lambda_2}{2}}\right\},$$

$\hat{I}_N(\lambda)$ 由(5.217)定义。

2. 在第 1 题的条件下，方差

$$D(\hat{I}_N(\lambda))=\begin{cases}4\sigma^4+O\left(\dfrac{1}{N^2}\right), & \lambda\neq 0,\ \pm\pi;\\ 8\sigma^4, & \lambda=0,\ \pm\pi.\end{cases}$$

特别对 $\lambda_p=\dfrac{2p\pi}{N}$,

$$D(\hat{I}_N(\lambda_p))=\begin{cases}4\sigma^4, & p\neq 0,\dfrac{N}{2}(N\text{偶});\\ 8\sigma^4, & p=0,\dfrac{N}{2}.\end{cases}$$

3. 在第 1 题条件下，对 $\lambda_1\neq\lambda_2,-\lambda_2$

$$\operatorname{Cov}(\hat{I}_N(\lambda_1),\ \hat{I}_N(\lambda_2))=\begin{cases}0, & \lambda_1=\dfrac{2p\pi}{N},\lambda_2=\dfrac{2q\pi}{N};\\ O\left(\dfrac{1}{N^2}\right), & \text{若 } |\lambda_1\pm\lambda_2|\gg\dfrac{2\pi}{N}.\end{cases}$$

4. 设谱窗核函数 $K(t)$ 除满足(5.165)对应的(a)条件外，还存在 $q>0$，使

$$\lim_{x\to 0}\frac{1-K(x)}{|x|^q}=h_q<+\infty,$$

并选 m_N 使 $m_N, Nm_N^{-1}, Nm_N^{-q}$ 都趋于无穷. 设 x_t 是实正态平稳列

共谱密度连续，

$$\sum_{n=-\infty}^{\infty}|n|^{q}|B_x(n)|<+\infty,$$

则

$$\hat{f}_N(\lambda)=\frac{1}{2\pi}\sum_{k=-m_N}^{m_N}\rho_k w_N(k)e^{-ik\lambda}$$

和 $f_x(\lambda)$ 的偏度有

$$\lim_{N\to\infty}m_N^q(E(\hat{f}_N(\lambda))-f_x(\lambda))=-\frac{h_q}{2\pi}\sum_{n=-\infty}^{\infty}|n|^{q}B_x(n)e^{-in\lambda}.$$

5. 试构造一串随机平稳列 $\{x_t\}$，它的周期图 $I_N(\lambda)$ 并不依概率收敛到谱密度 $f_x(\lambda)$。

第六章　多维时间序列介绍

§1　多维平稳序列

1°　多维平稳序列的定义，H_X 空间

在现实世界中，许多因素都是互相关联互相制约的，因此在研究随机过程时只研究一维的是不够的，这不仅是应用中需要有多维的模型，在理论上多维过程也有它独特的问题和兴趣。本章将简要地介绍这方面的若干主要结果。限于篇幅，有一些定理的证明将略去，有兴趣的读者可参看 Розанов[29]，Brillinger[12] 和 Hannan[82]三本有关多维时间序列的书籍。

定义6.1　设$\boldsymbol{X}_t=(x_1(t),x_2(t),\cdots,x_n(t))'$，$t\in T$ 是复值的 n 维随机序列，$E|x_k(t)|^2<+\infty$，$k=\overline{1,n}$。若对 $t\in T$ 有

$$
\begin{aligned}
&Ex_k(t)\equiv a_k, \qquad k=\overline{1,n},\\
&Ex_k(t+\tau)\overline{x_l(t)}=B_{kl}(\tau), \qquad\qquad (6.1)\\
&t,t+\tau\in T; \qquad k,l=\overline{1,n},
\end{aligned}
$$

则称 $\boldsymbol{X}_t$ 是 n 维平稳列，$x_k(t)$和 $x_l(t)$是平稳相关的。

若不声明则以下均假定 $E\boldsymbol{X}_t=0$，并称①

$$\boldsymbol{B}(\tau)=E\boldsymbol{X}_{t+\tau}\boldsymbol{X}_t^*=\boldsymbol{R}(\tau)=[Ex_k(t+\tau)\bar{x}_l(t)]_{k,l=\overline{1,n}} \qquad (6.2)$$

为 $\boldsymbol{X}_t$ 的相关矩阵(或协方差矩阵)。

定理6.1　设 $\boldsymbol{B}(\tau)$为 n 维平稳列的相关矩阵，则：

(a)　$\boldsymbol{B}(0)\geqslant 0$；

(b)　$\boldsymbol{B}^*(t_1-t_2)=\boldsymbol{B}(t_2-t_1)$；

(c)　$|B_{k,j}(\tau)|\leqslant\sqrt{B_{kk}(0)}\ \sqrt{B_{jj}(0)}$；

---

①　(6.2) 中的 $E\boldsymbol{X}_{t+\tau}\boldsymbol{X}_t^*$，* 表共轭转置，以后用"′"只代表转置。

(d) 对任意的正整数N和复向量 $\boldsymbol{Z}_k$($\boldsymbol{Z}_k$ 为 n 维复向量)，$t_k \in T$，$k=1,2,\cdots,N$，均有

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\boldsymbol{Z}_j^*\boldsymbol{B}(t_j-t_k)\boldsymbol{Z}_k\geqslant 0. \qquad (6.3)$$

证明 (a) 对任意的n维复向量$\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha}^*\boldsymbol{B}(0)\boldsymbol{\alpha}=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i B_{ij}(0)\overline{\alpha_j}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i EX_i(t)\overline{X_j(t)}\bar{\alpha}_j$$

$$=E\left|\sum_{i=1}^{n}\alpha_i x_i(t)\right|^2\geqslant 0.$$

(b) 因为

$$B_{kj}(t_1-t_2)=Ex_k(t_1)\overline{x_j(t_2)}$$
$$=\overline{Ex_j(t_2)\overline{x_k(t_1)}}=\overline{B_{jk}(t_2-t_1)}.$$

(c) 由 Schwartz 不等式可得。

(d) 记

$$\boldsymbol{Z}_k=(z_k^{(1)},z_k^{(2)},\cdots,z_k^{(n)})',\quad k=1,2,\cdots,N,$$

则

$$\boldsymbol{Z}_j^*\boldsymbol{B}(t_j-t_k)\boldsymbol{Z}_k=\sum_{s,t=1}^{n}\bar{z}_j^{(s)}b_{st}(t_j-t_k)z_k^{(t)}$$

$$=\sum_{s,t=1}^{n}\bar{z}_j^{(s)}E(x_s(t_j)\bar{x}_t(t_k))z_k^{(t)}$$

$$=E\left(\sum_{s=1}^{n}\bar{z}_j^{(s)}x_s(t_j)\right)\left(\overline{\sum_{t=1}^{n}\bar{z}_k^{(t)}\cdot x_t(t_k)}\right).$$

故

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\boldsymbol{Z}_j^*\boldsymbol{B}(t_j-t_k)\boldsymbol{Z}_k$$

$$
=E\left\{\sum_{j=1}^{N}\left(\sum_{s=1}^{n} z_j^{(s)} x_s(t_j)\right)\right\}\overline{\left\{\sum_{k=1}^{N}\left(\sum_{l=1}^{n} z_k^{(l)} x_l(t_k)\right)\right\}}
$$

$$
=E\left|\sum_{j=1}^{N}\sum_{s=1}^{n} z_j^{(s)} x_s(t_j)\right|^2 \geqslant 0. \quad \blacksquare
$$

和前几章一样，我们仍用 H 表示概率空间 $(\Omega,\mathscr{F},P)$ 上一切有二阶矩的随机变量所组成的 Hilbert 空间。以下我们将证明：一个 $n\times n$ 的矩阵 $\boldsymbol{B}(\tau)$，它是某个 n 维平稳列的相关函数阵的充分必要条件是 $\boldsymbol{B}(\tau)$ 是满足(6.3)的非负定函数阵。为此，我们先证

引理6.1 设 H 是 Hilbert 空间，M 是其中一个子集 $(M\subset H)$。在 M 上定义了等距算子 T，即对 $x,y\in M$，$(x,y)=(Tx,Ty)$。令 $H_M=\mathscr{L}\{M\}$（封闭的线性流型），则 T 可以把定义域扩张到 H_M，并且仍然保持等距，这种扩张不仅存在而且唯一。

证明 首先说明扩张是唯一的。设 T 扩张后为 T^*（即在 M 中 $T^*=T$）。对 $x\in H_M$，必有 $x_n\in M$ 使 $\lim\limits_{n} x_n=x$。于是

$$
\|T^*x-T^*x_n\|=\|x-x_n\|\to 0,\quad n\to\infty,
$$

即

$$
\lim_{n} T^*x_n=T^*x=\lim_{n} Tx_n.
$$

若又有 $\tilde{T}^*$ 亦是扩张，同理对 $x\in H_M$，

$$
\|\tilde{T}^*x-\tilde{T}^*x_n\|=\|\tilde{T}^*(x-x_n)\|=\|x-x_n\|\to 0,
$$

则

$$
\lim_{n}\tilde{T}^*x_n=\lim_{n}Tx_n=T^*x=\tilde{T}^*x.
$$

表明 T^* 和 $\tilde{T}^*$ 是相同的。

若 M 不是线性集，则首先将 M 扩充到线性集

$$
L(M)=\left\{\xi;\ \xi=\sum_{1}^{n}\alpha_i x_i,\ x_i\in M\right\}.
$$

定义

$$T\xi = \sum_{1}^{n} a_i T x_i. \qquad (6.4)$$

由

$$(T\xi, T\eta) = \sum_{1}^{n} \sum_{1}^{m} a_i \bar{\beta}_j (x_i, y_j) = (\xi, \eta) \qquad (6.5)$$

知T在$L(M)$上保持内积从而保距，并且(6.4)是一义的，而且是线性算子。

对$L(M)$可考虑其闭包$\mathscr{L}(M)$．对于$x_n \in L(M)$，因为$\lim_n x_n = x \in \mathscr{L}(M)$，由(6.5)，可知$\{Tx_n\}$为Cauchy列，从而存在$\xi$，$\xi = \lim_n Tx_n$．记

$$Tx = \zeta = \lim_n Tx_n, \qquad (6.6)$$

则T在$\mathscr{L}(M)$上是等距、线性算子．事实上

$$\begin{aligned}
T(\alpha x + \beta y) &= T(\alpha \lim_n x_n + \beta \lim_n y_n) \\
&= T(\lim_n(\alpha x_n + \beta y_n)) \\
&= \lim_n T(\alpha x_n + \beta y_n) \\
&= \lim_n(\alpha T x_n + \beta T y_n) \\
&= \alpha T x + \beta T y.
\end{aligned}$$

$$\begin{aligned}
(Tx, Ty) &= (\lim_n Tx_n, \lim_m Ty_n) \\
&= \lim_n \lim_m (x_n, y_m) = (x, y). \quad \blacksquare
\end{aligned}$$

引理6.2 设$X_t = (x_1(t), \cdots, x_n(t))'$为$n$维平稳列，记

$$H_X = \mathscr{L}\{x_i(t);\ i = \overline{1,n},\ t = 0, \pm 1, \cdots\}, \qquad (6.7)$$

并定义

$$M = \{x_k(t);\ k = \overline{1,n},\ t = 0, \pm 1, \cdots\}$$

上的酉算子U：

$$Ux_k(t) = x_k(t+1),\quad k = \overline{1,n},\ t = 0, \pm 1, \cdots, \qquad (6.8)$$

则U可唯一地扩张到H_X并且仍为酉算子。

证明 由引理6.1，(6.8)定义的U是M上的等距算子：

$$(Ux_i(t), Ux_j(s)) = (x_i(t+1), x_j(s+1)).$$
$$= (x_i(t), x_j(s)),$$

故U可唯一地扩张到$H_x = \mathscr{L}\{M\}$并且还是酉算子。 |

定理6.2 为了能够在H中找到一串点列$u_1, u_2, \cdots$使$(u_m, u_n) = c_{m,n}$，而$(c_{m_i, m_j})_{1 \leq i,j \leq k}$是给定的Hermite矩阵($k=1,2,\cdots$)，其充分必要条件是对任何$k, m_1, m_2 \cdots, m_k$，$(c_{m_i, m_j})_{1 \leq i,j \leq k}$是非负定的。

证明 由

$$0 \leqslant \left\| \sum_{i=1}^{k} a_i u_{m_i} \right\|^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} a_i \bar{a}_j c_{m_i m_j}$$

知必须性显然。以下证充分性，分二步证.

(a) 在H中，对有限值n，必存在一串点列$u_1^{(n)}, u_2^{(n)}, \cdots, u_n^{(n)}$，使

$$(u_i^{(n)}, u_j^{(n)}) = c_{ij} \quad (1 \leqslant i, j \leqslant n).$$

事实上，由于$(c_{ij})_{1 \leq i,j \leq n} = \boldsymbol{C} \geqslant 0$，设其秩为$m$，由Hermite阵性质必存在一个非奇异正交变换$\boldsymbol{A}$使

$$\begin{pmatrix} \boldsymbol{I}_m & \\ & 0 \end{pmatrix} = \boldsymbol{A}^* \boldsymbol{C} \boldsymbol{A} = \boldsymbol{A}^{-1} \boldsymbol{C} \boldsymbol{A}. \tag{6.9}$$

简记

$$\boldsymbol{A} = (a_{ij}^{(n)}), \quad \boldsymbol{A}^{-1} = (\bar{a}_{ji}^{(n)}), \tag{6.10}$$

则由

$$\boldsymbol{C} = \boldsymbol{A} \begin{pmatrix} \boldsymbol{I}_m & \\ & 0 \end{pmatrix} \boldsymbol{A}^{-1} \tag{6.11}$$

可得

$$c_{ik} = \sum_{j=1}^{m} a_{ij}^{(n)} \bar{a}_{kj}^{(n)}. \tag{6.12}$$

在H中可取m个标准正交点列

$$\boldsymbol{Z}^{(n)} = (z_1^{(n)}, \cdots, z_m^{(n)})',$$

令

$$u^{(n)}=A_m Z^{(n)}=(u_1^{(n)},u_2^{(n)},\cdots,u_n^{(n)})', \tag{6.13}$$

其中

$$A_m=A\begin{pmatrix} I_m \\ 0 \end{pmatrix}. \tag{6.14}$$

故

$$\begin{aligned}(u_i^{(n)},u_j^{(n)}) &= \left(\sum_{k=1}^{m} a_{ik}^{(n)} z_k^{(n)}, \sum_{l=1}^{m} a_{jl}^{(n)} z_l^{(n)}\right) \\ &= \sum_{k=1}^{m}\sum_{l=1}^{m} a_{ik}^{(n)} \bar{a}_{jl}^{(n)} \delta_{k,l} \\ &= \sum_{k=1}^{m} a_{ik}^{(n)} \bar{a}_{jk}^{(n)} = c_{ij}. \end{aligned} \tag{6.15}$$

由上看对任何有限的 $n\times n$ 阵 (c_{ij}) 均有 H 中的 $u^{(n)}$ 使(6.15)成立，令 $n=1,2,\cdots$ 遂得

$$(u_1^{(1)}),\ (u_1^{(2)},u_2^{(2)}),\cdots,(u_1^{(n)},u_2^{(n)},\cdots,u_n^{(n)}),\cdots. \tag{6.16}$$

(b) 用归纳法来选取满足定理要求的无穷序列。

对 $N=1$，令

$$u_1=u_1^{(1)}.$$

对任意的 N，设已选到 $(u_1,u_2,\cdots,u_N)$ 使得

$$(u_m,u_n)=c_{m,n}\ \ (1\leqslant m,n\leqslant N). \tag{6.17}$$

对 $N+1$，用以下方法来选第 $N+1$ 个元素。由 (6.16)已可找到

$$u^{(N+1)}=(u_1^{(N+1)},u_2^{(N+1)},\cdots,u_{N+1}^{(N+1)})',$$

于是定义等距算子 T_N：

$$T_N(u_k^{(N+1)})=u_k\ \ (1\leqslant k\leqslant N), \tag{6.18}$$

并记

$$M=\{u_1^{(N+1)},\cdots,u_N^{(N+1)}\},$$

则 T_N 是 M 上的等距算子：

$$(T_N u_i^{(N+1)},\ T_N u_j^{(N+1)})=(u_i,u_j)=c_{ij}=(u_i^{(N+1)},u_j^{(N+1)}).$$

由引理6.1，T_N 可以从M扩充到全空间H。这样，$u_{N+1}^{(N+1)}$ 亦有像 $T_N(u_{N+1}^{(N+1)})$，记为 u_{N+1}。由 T_N 之等距性

$$(u_{N+1},u_k)=(T_N u_{N+1}^{(N+1)},T_N u_k^{(N+1)})$$
$$=(u_{N+1}^{(N+1)},u_k^{(N+1)})=c_{N+1,k}.$$

这样即可得所求的点列 $u_1,u_2,\cdots,u_n,\cdots$。 ▌

定理6.2推论1 $n\times n$ 的矩阵函数 $\boldsymbol{R}(\tau)$是某个 n 维平稳列的相关函数的充分必要条件是 $\boldsymbol{R}(\tau)$为非负定的Hermite矩阵函数①。

证明 必需性由定理6.1可得，充分性可由定理6.2导出。首先，对平面上的部分格子点

$$\{(\mu,t)\,|\,\mu=1,2,\cdots,n;\ t=0,\pm1,\pm2\cdots\} \tag{6.19}$$

可排成一维形式的串

$$\{(\mu_1,t_1),\cdots,(\mu_q,t_q),\cdots\}②. \tag{6.19$'$}$$

记

$$c_{\mu_p\mu_q}=R_{\mu_p\mu_q}(t_p-t_q),\quad 1\leqslant p,q\leqslant m. \tag{6.20}$$

而

$$\boldsymbol{R}(\tau)=(R_{k,j}(\tau))_{n\times n},$$

于是$(c_{\mu_p\mu_q})_{m\times m}\geqslant0$，并且是Hermite的。由定理6.2，存在一串$H$中的点列 $x_{\mu_1}(t_1),x_{\mu_2}(t_2),\cdots$，使得

$$(x_{\mu_p}(t_p),x_{\mu_q}(t_q))=c_{\mu_p\mu_q}=R_{\mu_p\mu_q}(t_p-t_q). \tag{6.21}$$

再将$\{x_{\mu_i}(t_i),i=1,2,\cdots\}$按原来(6.19)的顺序写出即得

$$\{x_\mu(t)\,|\,\mu=1,2,\cdots,n,t=0,\pm1,\pm2,\cdots\}.$$ ▌

和一维的场合类似，我们也可给出 n 维 Gauss 存在定理。

定理6.2推论2 设 $\boldsymbol{R}(t)$是 $n\times n$ 的非负定Hermite矩阵函数，则必存在一个 n 维复 Gauss 过程 $\boldsymbol{X}_t$，以 $\boldsymbol{R}(t)$ 为其相关矩阵函数，均值为零。

类似于 Хинцин-Bochner 定理，多维场合下相关函数也有谱

① 即 (6.3) 式成立。

② 如$(1,0),\cdots,(n,0),(1,1),\cdots,(n,1),(1,-1),\cdots,(n,-1),\cdots$.

表示。

定理6.3 设 $x_1(t), x_2(t)$ 为平稳相关的序列，$R_{12}(\tau)$ 为相关函数，则

$$R_{12}(\tau) = \int_{-\pi}^{\pi} e^{i\tau\lambda} \, dF_{12}(\lambda), \tag{6.22}$$

其中 $F_{12}(\lambda)$ 为定义于 $-\pi \leqslant \lambda \leqslant \pi$ 上的复函数，左连续，有界变差并且 $F_{12}(-\pi)=0$，满足这一性质的 $F_{12}(\lambda)$ 是唯一的。

这一定理的证明可参看[4]，[29]。而 $F_{12}(\lambda)$ 可用以下有界变差函数的 Fourier 变换得到：

$$\begin{cases} F_{12}(\lambda) = \dfrac{1}{2\pi} W_{12}(\lambda+0) + c; & (6.23) \\ W_{12}(\lambda) = R_{12}(0)\lambda - \displaystyle\sum_{k \neq 0} \dfrac{R_{12}(k)}{ik} e^{-ik\lambda}, & (6.24) \end{cases}$$

其中常数 c 由 $F_{12}(-\pi)=0$ 确定。一般说来 $F_{12}(\lambda)$ 是复函数

$$F_{12}(\lambda) = \overline{F_{21}(\lambda)}, \tag{6.25}$$

即

$$R_{12}(k) = \overline{R_{21}(-k)}$$

推论 设 $\boldsymbol{X}_t$ 是 n 维平稳列，$\boldsymbol{R}(\tau)$ 为其相关矩阵函数，则必存在一个 $n \times n$ 的矩阵函数

$$\boldsymbol{F}(\lambda) = (F_{kj}(\lambda))_{n \times n} \tag{6.26}$$

使得

$$\boldsymbol{R}(\tau) = \int_{\Pi} e^{i\lambda\tau} \, d\boldsymbol{F}(\lambda), \quad \tau = 0, \pm 1, \cdots, \tag{6.27}$$

其中 $F_{kj}(\lambda)$ 是定义于 Π 上的左连续复有界变差函数，并且 $F_{kj}(-\pi)=0$，这样谱阵 $\boldsymbol{F}(\lambda)$ 还是唯一的。

若对 $\Delta \in \mathscr{B}(\Pi)$，记

$$F_{kj}(\Delta) = \int_{\Delta} dF_{kj}(\lambda). \tag{6.28}$$

则

$$\boldsymbol{F}(\Delta) = (F_{k,j}(\Delta))_{1 \leqslant k,j \leqslant n} \geqslant 0. \tag{6.29}$$

若 $\Delta_k \in \mathscr{B}(\Pi)$，互不相交，必有

$$F\left(\bigcup_{k=1}^{\infty} \Delta_k\right) = \sum_{k=1}^{\infty} F(\Delta_k). \tag{6.30}$$

可见，$\mathrm{d}\boldsymbol{F}(\lambda)$ 为矩阵测度。

如果 $\mathrm{d}F_{k,j}(\lambda) \ll \mathrm{d}\lambda$，则必有

$$\frac{\mathrm{d}\boldsymbol{F}(\lambda)}{\mathrm{d}\lambda} = \boldsymbol{f}(\lambda) = \left(\frac{\mathrm{d}F_{k,j}(\lambda)}{\mathrm{d}\lambda}\right)_{1 \leqslant k, j \leqslant n} \tag{6.31}$$

存在，称为 $\boldsymbol{X}_t$ 的谱密度矩阵。

一个通用的保证 $\boldsymbol{f}(\lambda)$ 存在的充分条件为 $\boldsymbol{X}_t$ 的相关阵的元素 $R_{kj}(\tau)$ 满足

$$\sum_{\tau=-\infty}^{\infty} |R_{kj}(\tau)| < +\infty \tag{6.32}$$

对一切 $1 \leqslant k, j \leqslant n$ 成立。

2° 多维平稳序列的例子

例6.1（n 维白噪声）　设 $\boldsymbol{\varepsilon}_t = (\varepsilon_1(t), \cdots, \varepsilon_n(t))'$ 是 n 维随机序列，满足

(a) $E\varepsilon_k(t) \equiv 0$,　$k = \overline{1, n}$,　$t \in 0, \pm 1, \pm 2, \cdots$;

(b) $E\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^* = \delta_{t,s} \boldsymbol{\Sigma}_{n \times n}$, $\tag{6.33}$

其中 $\boldsymbol{\Sigma}_{n \times n}$ 为正定阵。特别当 $\boldsymbol{\Sigma}_{n \times n} = \boldsymbol{I}_n$（单位阵）时则称为是标准白噪声列。

需注意的是由(6.33)的定义表明对同一时刻 $\boldsymbol{\varepsilon}_t$ 各分量未必是相互独立的。

例6.2（MA($+\infty$)线性序列）　设 $\boldsymbol{\varepsilon}_t$ 是 n 维白噪声标准列，$\boldsymbol{A}_k$，$k = 0, 1, 2, \cdots$ 是一串 $n \times n$ 的矩阵，满足

$$\sum_{k=0}^{\infty} \operatorname{tr}(\boldsymbol{A}_k \boldsymbol{A}_k^*) < +\infty, \tag{6.34}$$

测在 H 空间中存在

$$\zeta_t = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k} \quad (L^2) \tag{6.35}$$

并且 ζ_t 是平稳列。

需说明的是：在H中，对随机向量$\zeta=(\zeta_1,\zeta_2,\cdots,\zeta_n)$定义

$$\|\zeta\| = \sqrt{\sum_{l=1}^{n} E|\zeta_l|^2}, \tag{6.36}$$

称向量级数$\sum_{k=0}^{\infty} \eta_k$在$H$中是$L^2$收敛的，假若存在随机变量$\zeta\in H$使

$$\lim_{N\to\infty} \left\|\sum_{k=0}^{N} \eta_k - \zeta\right\| = 0. \tag{6.37}$$

容易看出

$$\sum_{k=0}^{\infty} \eta_k < +\infty \Longleftrightarrow \left\|\sum_{N}^{M} \eta_k\right\| \to 0, \quad N, M\to\infty. \tag{6.38}$$

本例(6.34)正好保证了(6.35)的收敛性。

记

$$A_k = \begin{pmatrix} A_{k,1} \\ A_{k,2} \\ \vdots \\ A_{k,n} \end{pmatrix}, \quad k=0,1,2,\cdots,$$

则

$$\begin{aligned}
&\left\|\sum_{k=N}^{N'} A_k \varepsilon_{t-k}\right\|^2 \\
&= \sum_{l=1}^{n} E\left|\sum_{k=N}^{N'} A_{k,l}\varepsilon_{t-k}\right|^2 \\
&= \sum_{l=1}^{n} \sum_{k,k'=N}^{N'} A_{k,l} E(\varepsilon_{t-k}\varepsilon_{t-k'}^{*})(A_{k',l})^{*}
\end{aligned}$$

$$= \sum_{l=1}^{n} \sum_{k,k'=N}^{N'} A_{k,l} I \delta_{k,k'} (A_{k',l})^*$$

$$= \sum_{l=1}^{n} \sum_{k=N}^{N'} \sum_{j=1}^{n} |a_{k,l}^{(j)}|^2 = \sum_{k=N}^{N'} \mathrm{tr}(\boldsymbol{A}_k \boldsymbol{A}_k^*)$$

$$\to 0 \quad (N, N' \to \infty). \tag{6.39}$$

由此保证了级数向量的收敛性。

若记 $\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^*) = \|\boldsymbol{A}\|^2$，则

$$\sum_k \boldsymbol{A}_k \boldsymbol{\varepsilon}_{t-k} \text{级数收敛} \Longleftrightarrow \sum_k \|\boldsymbol{A}_k\|^2 < +\infty. \tag{6.40}$$

$$\boldsymbol{R}_{\zeta\zeta}(m) = E\boldsymbol{\zeta}_{m+s}\boldsymbol{\zeta}_s^*$$

$$= \sum_{k,l=0}^{\infty} \boldsymbol{A}_k E \boldsymbol{\varepsilon}_{m+s-k} \boldsymbol{\varepsilon}_{s-l}^* \boldsymbol{A}_l^*$$

$$= \sum_{l=0}^{\infty} \boldsymbol{A}_{l+m} \boldsymbol{A}_l^*. \tag{6.41}$$

例6.3（马氏平稳列） 设 $\boldsymbol{X}_t$ 是 p 维平稳列，具有马氏性（参看第二章§4）

$$E[\boldsymbol{X}_{n+s} | H_{\boldsymbol{X}}(n)] = E[\boldsymbol{X}_{n+s} | \mathscr{L}_{\boldsymbol{X}}(n)]^{①}, \tag{6.42}$$

其中

$$H_{\boldsymbol{X}}(n) = \mathscr{L}\{x_i(t);\ t \leqslant n,\ i = \overline{1,p}\},$$

$$\mathscr{L}_{\boldsymbol{X}}(n) = \mathscr{L}\{x_i(n);\ i = \overline{1,p}\}.$$

假定 $\boldsymbol{R}_{\boldsymbol{X}}(0) = E\boldsymbol{X}(n)\boldsymbol{X}^*(n)$ 非退化，则

$$\boldsymbol{R}_{\boldsymbol{X}}(m) = \boldsymbol{A}^m \boldsymbol{R}_{\boldsymbol{X}}(0), \tag{6.43}$$

其中

$$\boldsymbol{A} = \boldsymbol{R}_{\boldsymbol{X}}(1) \boldsymbol{R}_{\boldsymbol{X}}^{-1}(0). \tag{6.44}$$

证明 由(6.42)可得

$$E[\boldsymbol{X}_{n+1} - \boldsymbol{A}\boldsymbol{X}_n]\boldsymbol{X}_n^* = \boldsymbol{R}_{\boldsymbol{X}}(1) - \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{X}}(0) = 0$$

---

① $E[y|H]$ 表示 y 在 H 上投影。

由此可得(6.44)，并且 $\boldsymbol{X}_n$ 与 $\boldsymbol{\eta}_n=\boldsymbol{X}_{n+1}-\boldsymbol{A}\boldsymbol{X}_n$ 是正交的。于是

$$E[\boldsymbol{X}_{n+1}|\mathscr{L}_X(n)]=E[\boldsymbol{A}\boldsymbol{X}_n+\boldsymbol{\eta}_n|\mathscr{L}_X(n)]$$
$$=\boldsymbol{A}E[\boldsymbol{X}_n|\mathscr{L}_X(n)]=\boldsymbol{A}\boldsymbol{X}_n. \tag{6.45}$$

利用如下事实：当 $H_1\subset H_2\subset H_X$，则

$$E[E[\boldsymbol{X}|H_1]|H_2]=E[\boldsymbol{X}|H_1],$$
$$E[E[\boldsymbol{X}|H_2]|H_1]=E[\boldsymbol{X}|H_1]. \tag{6.46}$$

对 $n,s\geqslant 1$，利用马氏性

$$\begin{aligned}
&E[\boldsymbol{X}_{n+s}|H_X(s)]\\
&=E[\boldsymbol{X}_{n+s}|\mathscr{L}_X(s)]\\
&=E[E[\boldsymbol{X}_{n+s}|H_X(n+s-1)]\mathscr{L}_X(s)]\\
&\quad(\because H_X(n+s-1)\supset\mathscr{L}_X(n))\\
&=E[E[\boldsymbol{X}_{n+s}|\mathscr{L}_X(n+s-1)]|\mathscr{L}_X(s)] &(6.47)\\
&=E[\boldsymbol{A}\boldsymbol{X}_{n+s-1}|\mathscr{L}_X(s)]=\boldsymbol{A}E[\boldsymbol{X}_{n+s-1}|\mathscr{L}_X(s)] &(6.48)\\
&=\boldsymbol{A}^2E[\boldsymbol{X}_{n+s-2}|\mathscr{L}_X(s)]=\cdots\\
&=\boldsymbol{A}^nE[\boldsymbol{X}_s|\mathscr{L}_X(s)]=\boldsymbol{A}^n\boldsymbol{X}_s. &(6.49)
\end{aligned}$$

对 $n\geqslant 1$ 以上表明

$$\boldsymbol{X}_{n+s}-\boldsymbol{A}^n\boldsymbol{X}_s=\boldsymbol{\Delta}_{n,s}\perp H_X(s).$$

可见

$$E\boldsymbol{X}_{n+s}\boldsymbol{X}_s^*=\boldsymbol{A}^nE[\boldsymbol{X}_s\boldsymbol{X}_s^*]+E[\boldsymbol{\Delta}_{n,s}\boldsymbol{X}_s^*]$$

即

$$\boldsymbol{R}_X(n)=\boldsymbol{A}^n\boldsymbol{R}_X(0). \quad \blacksquare$$

3° H_X 与 $L^2(\mathrm{d}\boldsymbol{F})$ 的同构对应

以下假定 n 维平稳列 $\boldsymbol{X}_t$ 是有谱密度阵 $\boldsymbol{f}(\lambda)$①。考虑函数向量空间 $L^2(\mathrm{d}\boldsymbol{F})$：

$$L^2(\mathrm{d}\boldsymbol{F})=\{\underset{\smile}{\varphi}=(\varphi_1,\varphi_2,\cdots,\varphi_n);\ \int_\Pi\underset{\smile}{\varphi}\boldsymbol{f}(\lambda)\underset{\smile}{\varphi}^*\mathrm{d}\lambda<+\infty\}, \tag{6.50}$$

① 若无密度，则内积定义为

$$(\underset{\smile}{\varphi},\underset{\smile}{\psi})=\int_\Pi\underset{\smile}{\varphi}\mathrm{d}F(\lambda)\underset{\smile}{\psi}^*. \tag{6.51}$$

其中定义内积

$$(\underset{\smile}{\varphi},\underset{\smile}{\psi}) = \int_{\Pi} \underset{\smile}{\varphi} \boldsymbol{f}(\lambda) \underset{\smile}{\psi}^* \mathrm{d}\lambda, \quad \underset{\smile}{\varphi},\underset{\smile}{\psi} \in L^2(\mathrm{d}\boldsymbol{F}). \tag{6.51}$$

在 $L^2(\mathrm{d}\boldsymbol{F})$ 中 称 $\underset{\smile}{\varphi} = \underset{\smile}{\psi}$ 是指

$$\|\underset{\smile}{\varphi} - \underset{\smile}{\psi}\|^2 = \int_{\Pi} (\underset{\smile}{\varphi} - \underset{\smile}{\psi}) \boldsymbol{f}(\lambda) (\underset{\smile}{\varphi} - \underset{\smile}{\psi})^* \mathrm{d}\lambda = 0. \tag{6.52}$$

因而未必有 $\underset{\smile}{\varphi} = \underset{\smile}{\psi}[\mathrm{d}\lambda]$。和一维序列类似我们可建立

$$H_X = \mathscr{L}\{x_i(t);\ i = \overline{1,n},\ t = 0, \pm 1, \cdots\}$$

和 $L^2(\mathrm{d}\boldsymbol{F})$ 的对应关系：

定理6.4 设 $\boldsymbol{X}_t$ 是平稳列，有谱密度阵 $\boldsymbol{f}(\lambda)$，则存在一个映射 Ψ，使得在此映射下，H_X 与 $L^2(\mathrm{d}\boldsymbol{F})$ 是同构的 Hilbert 空间，并且 Ψ 将 $\boldsymbol{X}_t$ 的分量 $x_k(t)$ 映射为 $\underset{\smile}{\delta}_k e^{i\lambda t}$，其中

$$\underset{\smile}{\delta}_k = (\overbrace{0, \cdots, 1}^{k}, 0, \cdots 0). \tag{6.53}$$

定理6.5 设 $\boldsymbol{X}_t$ 是 n 维平稳列，则 $\boldsymbol{X}_t$ 可表为

$$\boldsymbol{X}_t = \int_{\Pi} e^{i\lambda t}\, \mathrm{d}\boldsymbol{Z}(\lambda), \qquad t = 0, \pm 1, \cdots, \tag{6.54}$$

其中 $\mathrm{d}\boldsymbol{Z}(\lambda) = (\mathrm{d}z_1(\lambda), \cdots, \mathrm{d}z_n(\lambda))'$ 是 Π 上的 n 维正交随机测度，使

$$E z_k(\Delta_1)\overline{z_l(\Delta_2)} = \int_{\Delta_1 \cap \Delta_2} \mathrm{d}F_{kl}(\lambda),$$

$$\Delta_1, \Delta_2 \in \mathscr{B}(\Pi), \quad k, l = \overline{1,n}. \tag{6.55}$$

此中 $(F_{kl}(\lambda))_{n\times n}$ 是 $\boldsymbol{X}_t$ 的谱函数矩阵。

此定理的证明可参看[29]。我们只作以下的注释：

(1) 设 $\varphi_k(\lambda) \in L^2(\mathrm{d}F_{kk})$，$k = \overline{1,n}$，则

$$\underset{\smile}{\phi} = (\varphi_1, \varphi_2, \cdots, \varphi_n) \in L^2(\mathrm{d}\boldsymbol{F}).$$

事实上，由

$$\left|\int_{\Pi} \underset{\smile}{\varphi} \mathrm{d}\boldsymbol{F}\, \underset{\smile}{\varphi}^*\right| = \left|\sum_{k,l} \int_{\Pi} \varphi_k \overline{\varphi}_l\, f_{kl}(\lambda) \mathrm{d}\lambda\right|$$

$$\leqslant \sum_{k,l} \int_{\Pi} (|\varphi_k| \sqrt{f_{kk}(\lambda)}) (|\varphi_l| \sqrt{f_{ll}(\lambda)}) \mathrm{d}\lambda$$

$$\leqslant \sum_{k,l} \sqrt{\int_{\Pi} |\varphi_k|^2 f_{kk}(\lambda) \,\mathrm{d}\lambda} \sqrt{\int_{\Pi} |\varphi_l|^2 f_{ll}(\lambda) \mathrm{d}\lambda} < +\infty \tag{6.56}$$

知以上结论为真。但反之不对，这表明一般

$$\eta \in H_X = \mathscr{L}\{x_i(t);\ i=\overline{1,n},\ t=0,\pm 1,\cdots\}$$

并不一定能将它表为

$$\eta = \sum_{i=1}^{n} \eta_i,\quad \eta_i \in H_{x_i},\quad i=1,2,\cdots,n.$$

在[79]中江泽培给出如下的重要结果：

定理6.6 设 X_t 为 n 维平稳列，其谱密度阵 $f(\lambda)$ 几乎处处非退化，则 $\underset{\smile}{\varphi}=(\varphi_1,\varphi_2,\cdots,\varphi_n) \in L^2(\mathrm{d}F)$ 等价于 $\varphi_k \in L^2(\mathrm{d}F_{kk})$，$i=\overline{1,n}$ 的充分必要条件是：存在 $\varepsilon_0>0$ 使得

$$\frac{\det(f(\lambda))}{\prod_{k=1}^{n} f_{kk}(\lambda)} \geqslant \varepsilon_0 > 0 \quad \text{a.e.}[\mathrm{d}\lambda]. \tag{6.57}$$

(2) 集合

$$\Phi = \{\underset{\smile}{\varphi}=(\varphi_1,\varphi_2,\cdots,\varphi_n);\ \varphi_k \in L^2(\mathrm{d}F_{kk}),\ k=\overline{1,n}\} \tag{6.58}$$

在 $L^2(\mathrm{d}F)$ 中是稠密的。

事实上，令 $\Psi(\Phi)=\hat{H}$，不难看出 $\hat{H}$ 在 H_X 中是稠密的。

(3) 关于随机积分的定义。

对于 $\underset{\smile}{\varphi} \in \Phi$，对其分量已在一维讨论中定义了随机积分，因而定义

$$\int_{\Pi} \underset{\smile}{\varphi} \mathrm{d}Z(\lambda) = \sum_{k=1}^{n} \int_{\Pi} \varphi_k(\lambda) \mathrm{d}z_k(\lambda) \in \hat{H}. \tag{6.59}$$

对于 $\underset{\smile}{\varphi} \in L^2(\mathrm{d}F)$，由 Φ 在 $L^2(\mathrm{d}F)$ 中稠密，可选一串 $\underset{\smile}{\varphi}^{(N)} \in \Phi$，$\underset{\smile}{\varphi}^{(N)} \xrightarrow{L^2} \underset{\smile}{\varphi}$。记

$$\zeta^{(N)} = \sum_{k=1}^{n} \int_{\Pi} \varphi_k^{(N)} \mathrm{d} z_k ,$$

则

$$\|\zeta^{(N)} - \zeta^{(M)}\|^2 = \|\underset{\smile}{\varphi}^{(N)} - \underset{\smile}{\varphi}^{(M)}\|^2 \to 0, \quad N, M \to \infty .$$

表明必存在 $\zeta \in H_X$，使 $\zeta^{(N)} \to \zeta$。于是定义

$$\zeta = \int_{\Pi} \underset{\smile}{\varphi}(\lambda) \mathrm{d} \boldsymbol{Z}(\lambda) . \tag{6.60}$$

不难证明这样的对应是保内积的：

$$\begin{aligned}(\xi, \eta) &= \left(\int_{\Pi} \underset{\smile}{\varphi} \mathrm{d} \boldsymbol{Z}(\lambda), \int_{\Pi} \underset{\smile}{\psi} \mathrm{d} \boldsymbol{Z}(\lambda)\right) \\ &= \int_{\Pi} \underset{\smile}{\varphi} \, \mathrm{d} \boldsymbol{F} \underset{\smile}{\psi}^* = (\underset{\smile}{\varphi}, \underset{\smile}{\psi}) .\end{aligned} \tag{6.61}$$

(4) 记 $H_k(t) = \mathscr{L}\{x_k(s);\ s \leqslant t\}$，$k = \overline{1, n}$，则

$$H_k(t) \subset H_X(t) \subset H_X .$$

若取

$$\underset{\smile}{\varphi} = (\overbrace{0, \cdots, 0, \varphi_k}^{k}, 0 \cdots 0), \quad \varphi_k \in L^2(\mathrm{d} F_{kk}),$$

$$\underset{\smile}{\psi} = (\overbrace{0, \cdots, 0, \psi_l}^{l}, 0 \cdots 0), \quad \psi_l \in L^2(\mathrm{d} F_{ll}),$$

则

$$\left(\int_{\Pi} \underset{\smile}{\varphi} \mathrm{d} \boldsymbol{Z}(\lambda), \int_{\Pi} \underset{\smile}{\psi} \mathrm{d} \boldsymbol{Z}(\lambda)\right) = \int_{\Pi} \varphi_k \bar{\psi}_l \mathrm{d} F_{kl} = \left(\int_{\Pi} \varphi_k \mathrm{d} z_k, \int_{\Pi} \psi_l \mathrm{d} z_l\right). \tag{6.62}$$

§2 多维 ARMA 模型

1° 多维 ARMA 序列的定义

定义6.2 设 $\boldsymbol{X}_t$ 为 n 维平稳列，称它是多维 ARMA(p, q)，

假若它满足如下的多维随机差分方程

$$\begin{aligned}&\boldsymbol{X}_t+\boldsymbol{A}_1\boldsymbol{X}_{t-1}+\cdots+\boldsymbol{A}_p\boldsymbol{X}_{t-p}\\&=\boldsymbol{\varepsilon}_t+\boldsymbol{B}_1\boldsymbol{\varepsilon}_t+\cdots+\boldsymbol{B}_q\boldsymbol{\varepsilon}_{t-q},\end{aligned}\tag{6.62'}$$

其中 $\boldsymbol{B}_k,\boldsymbol{A}_l$ 皆为 $n\times n$ 的实方阵，并且

$$\begin{aligned}\boldsymbol{B}(z)&=\sum_{k=0}^{q}\boldsymbol{B}_k z^k,\quad \boldsymbol{B}_0=\boldsymbol{I},\\ \boldsymbol{A}(z)&=\sum_{k=0}^{p}\boldsymbol{A}_k z^k,\quad \boldsymbol{A}_0=\boldsymbol{I}\end{aligned}\tag{6.63}$$

的行列式 $\det\boldsymbol{B}(z),\det\boldsymbol{A}(z)$的根皆在单位圆外，并且 $\boldsymbol{A}(z),\boldsymbol{B}(z)$ 是左既约的，即若有

$$\boldsymbol{A}(z)=\boldsymbol{P}(z)\boldsymbol{A}_1(z),$$
$$\boldsymbol{B}(z)=\boldsymbol{P}(z)\boldsymbol{B}_1(z),$$

则 $\det\boldsymbol{P}(z)\equiv1$，并且 $\operatorname{rank}(\boldsymbol{A}_p;\boldsymbol{B}_q)=n$。而 $\boldsymbol{\varepsilon}_t$ 是 n 维白噪声列

$$E\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s^*=\delta_{t,s}\boldsymbol{D}_0,\tag{6.64}$$

而 $\boldsymbol{D}_0$ 非退化。

定义6.3 多维 ARMA 模型中，若 $p=0$，则称为 MA(q)模型；若 $q=0$ 称为 AR(p)模型。

例6.4 设 ε_t 为 2 维标准白噪声列，$\boldsymbol{X}_t$ 为满足

$$\boldsymbol{X}_t=\varepsilon_t-\begin{pmatrix}\frac{1}{3}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{3}\end{pmatrix}\varepsilon_{t-1}\tag{6.65}$$

方程的平稳列，则 $\boldsymbol{X}_t$ 为 2 维 MA(1)模型。此时

$$\begin{aligned}\boldsymbol{B}(z)&=\boldsymbol{I}_2-\begin{pmatrix}\frac{1}{3}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{3}\end{pmatrix}z\\&=\begin{pmatrix}1-\frac{1}{3}z&-\frac{1}{2}z\\-\frac{1}{2}z&1-\frac{1}{3}z\end{pmatrix},\end{aligned}\tag{6.66}$$

从而

$$\det \boldsymbol{B}(z) = \left(1-\frac{1}{3}z\right)^2 - \frac{1}{4}z^2$$

$$= 1-\frac{2}{3}z-\frac{5}{36}z^2$$

$$= -\frac{1}{36}(5z-6)(z+6) \qquad (6.67)$$

有二个根$z_1=\frac{6}{5}, z_2=-6$，它们皆在单位圆外。

例6.5 Grenander 和 Rosenblatt 在[68]中给出 如下 经济模型：

设 $x_1(t)$表 t 时的价格，$x_2(t)$表 t 时的供给量，则有如下经济模型成立

$$\begin{cases} x_1(t) = -bx_2(t) + \eta_1(t), \\ x_2(t) = ax_1(t-1) + \eta_2(t), \end{cases} \quad a, b > 0, \qquad (6.68)$$

$\eta_1(t), \eta_2(t)$为标准白噪声。若视为多维模型它是：

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -a & 0 \end{pmatrix}\begin{pmatrix} x_1(t-1) \\ x_2(t-1) \end{pmatrix} = \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \end{pmatrix}$$

或

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} ab & 0 \\ -a & 0 \end{pmatrix}\begin{pmatrix} x_1(t-1) \\ x_2(t-1) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix}. \qquad (6.69)$$

而

$$\boldsymbol{A}(z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} abz & 0 \\ -az & 0 \end{pmatrix} = \begin{pmatrix} 1+abz & 0 \\ -az & 1 \end{pmatrix},$$

从而要 $\det \boldsymbol{A}(z) = 1 + abz$ 的根 $z = -1/ab$ 在单位圆外就应 有 $ab < 1$。

$$\begin{cases} \varepsilon_1(t) = \eta_1(t) - b\eta_2(t), \\ \varepsilon_2(t) = \eta_2(t), \end{cases} \qquad (6.70)$$

$$D_0 = E\begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix}(\varepsilon_1(t), \varepsilon_2(t)) = \begin{pmatrix} 1+b^2 & -b \\ -b & 1 \end{pmatrix}. \quad (6.71)$$

可见(6.69)为2维AR(2)模型。

2° ARMA模型的相关函数阵和谱密度阵

和一维相仿，以下给出 n 维 AR(p)，MA(q)和ARMA(p,q)相关矩阵函数所应满足的差分方程和谱密度阵的参数表示。

定理6.6 设 X_t 为 n 维 AR(p)模型平稳列，模型方程为

$$A(U)X_t = \varepsilon_t, \quad (6.72)$$

则 $R_X(\tau)$ 相关阵函数满足以下 Yule-Walker 方程

$$A(U)R_X(m) = \begin{cases} D_0, & m = 0; \\ 0, & m > 1, \end{cases} \quad (6.73)$$

且其谱密度阵为

$$f(\lambda) = \frac{1}{2\pi}A^{-1}(e^{-i\lambda})D_0(A^{-1}(e^{-i\lambda}))^*, \quad (6.74)$$

其中

$$A(e^{-i\lambda}) = \sum_{k=0}^{p} A_k e^{-ik\lambda}, \quad (6.75)$$

$$A(U)R_X(m) = R_X(m) + A_1R_X(m-1) + \cdots + A_pR_X(m-p). \quad (6.76)$$

我们假定 X_t 有 Wold 分解式①

$$X_t = \sum_{k=0}^{\infty} C_k\varepsilon_{t-k}, \quad (6.77)$$

来看定理6.6，其中

$$E\varepsilon_t\varepsilon_s' = \delta_{t,s}I_n.$$

方程

① 在 ARMA 模型下多维平稳列 X_t 必有(6.77)的 Wold 展式，见[29]。

$$\sum_{k=0}^{p} \boldsymbol{A}_k \boldsymbol{X}_{t+m-k} = \boldsymbol{\varepsilon}_{t+m} = \boldsymbol{C}_0 \boldsymbol{\varepsilon}_{t+m} \tag{6.78}$$

两边乘 $\boldsymbol{X}_t^*$ 并取期望：

$$\sum_{k=0}^{p} \boldsymbol{A}_k E(\boldsymbol{X}_{t+m-k}\boldsymbol{X}'_t) = \sum_{k=0}^{\infty} \boldsymbol{C}_0 E\boldsymbol{\varepsilon}_{t+m}\boldsymbol{\varepsilon}'_{t-k}\boldsymbol{C}'_k,$$

$$\sum_{k=0}^{p} \boldsymbol{A}_k \boldsymbol{R}_X(m-k) = \sum_{k=0}^{\infty} \delta_{m,-k} \boldsymbol{C}_0 \boldsymbol{I} \boldsymbol{C}'_k$$

$$= \begin{cases} \boldsymbol{C}_0\boldsymbol{C}'_0 = \boldsymbol{D}_0, & m = 0; \\ 0, & m > 1. \end{cases} \tag{6.79}$$

又由(6.78)知

$$\sum_{k=0}^{p} \boldsymbol{A}_k \int_\Pi e^{i(t-k)\lambda}\, d\boldsymbol{Z}_X = \int_\Pi e^{it\lambda} d\boldsymbol{Z}_\varepsilon$$

或

$$\sum_{k=0}^{p} \boldsymbol{A}_k e^{-ik\lambda} d\boldsymbol{Z}_X = d\boldsymbol{Z}_\varepsilon,$$

即

$$d\boldsymbol{Z}_X = \boldsymbol{A}^{-1}(e^{-i\lambda}) d\boldsymbol{Z}_\varepsilon. \tag{6.80}$$

从而

$$\boldsymbol{X}_t = \int_\Pi e^{i\lambda t} \boldsymbol{A}^{-1}(e^{-i\lambda}) d\boldsymbol{Z}_\varepsilon,$$

$$\boldsymbol{R}_X(\tau) = \left(\int_\Pi e^{i\lambda(t+\tau)} \boldsymbol{A}^{-1}(e^{-i\lambda}) d\boldsymbol{Z}_\varepsilon,\ \int_\Pi e^{i\lambda t} \boldsymbol{A}^{-1}(e^{-i\lambda}) d\boldsymbol{Z}_\varepsilon\right)$$

$$= \int_\Pi e^{i\lambda\tau} \boldsymbol{A}^{-1}(e^{-i\lambda}) \boldsymbol{f}_\varepsilon(\lambda) (\boldsymbol{A}^{-1}(e^{-i\lambda}))^* d\lambda. \tag{6.81}$$

由谱函数的唯一性

$$\boldsymbol{f}_X(\lambda) = \frac{1}{2\pi}(\boldsymbol{A}^{-1}(e^{-i\lambda}))\boldsymbol{D}_0(\boldsymbol{A}^{-1}(e^{-i\lambda}))^*. \tag{6.82}$$

用以上类似的讨论方法可得：

定理6.7 设 $\boldsymbol{X}_t$ 为 n 维的 MA(q)序列，其模型方程为

$$\boldsymbol{X}_t=\sum_{k=0}^{q}\boldsymbol{B}_k\boldsymbol{\varepsilon}_{t-k}, \qquad \boldsymbol{B}_0=\boldsymbol{I}_n, \tag{6.83}$$

则其相关矩阵函数满足

$$\boldsymbol{R}_X(m)=\begin{cases}\sum\limits_{k=0}^{q-m}\boldsymbol{B}_{m+k}\boldsymbol{D}_0\boldsymbol{B}'_k, & 0\leqslant m\leqslant q;\\ 0, & m>q,\end{cases} \tag{6.84}$$

并且谱密度阵为

$$\boldsymbol{f}_X(\lambda)=\frac{1}{2\pi}\boldsymbol{B}(\mathrm{e}^{-i\lambda})\boldsymbol{D}_0\boldsymbol{B}^*(\mathrm{e}^{-i\lambda}). \tag{6.85}$$

定理6.8 设 $\boldsymbol{X}_t$ 为 n 维 ARMA(p,q)模型，其方程为

$$\boldsymbol{A}(U)\boldsymbol{X}_t=\boldsymbol{B}(U)\boldsymbol{\varepsilon}_t,$$

则其相关阵 $\boldsymbol{R}_X(\tau)$满足以下的差分方程

$$\boldsymbol{A}(U)\boldsymbol{R}_X(m)=0, \qquad \text{当 } m>q, \tag{6.86}$$

并且谱密度阵为

$$\boldsymbol{f}_X(\lambda)=\frac{1}{2\pi}(\boldsymbol{A}^{-1}(\mathrm{e}^{-i\lambda}))\boldsymbol{B}(\mathrm{e}^{-i\lambda})\boldsymbol{D}_0(\boldsymbol{B}(\mathrm{e}^{-i\lambda}))^* \cdot (\boldsymbol{A}^{-1}(\mathrm{e}^{-i\lambda}))^*. \tag{6.87}$$

3° 多维 Yule-Walker 方程的递推算法

类似于一维的场合，以下介绍如何用递推方法来解多维的 Yule-Walker 方程(6.73)，此法多称为 Levinson 递推。

若记

$$\begin{aligned}-\boldsymbol{A}_j^{(p)}&=\boldsymbol{A}_j, \qquad j=1,2,\cdots,p,\\ \mathscr{A}_j^{(p)}&=(\boldsymbol{A}_1^{(j)},\cdots,\boldsymbol{A}_j^{(p)})',\\ \mathscr{A}_p^{(p)}&=(\boldsymbol{A}_1^{(p)},\cdots,\boldsymbol{A}_p^{(p)})',\\ \boldsymbol{\eta}_p&=(\boldsymbol{R}(1),\cdots,\boldsymbol{R}(p))',\end{aligned} \tag{6.88}$$

则(6.73)相应于(2.100)是

$$\boldsymbol{\Gamma}_p \mathscr{A}_p^{(p)} = \boldsymbol{\eta}_p, \tag{6.89}$$

其中

$$\boldsymbol{\Gamma}_p = \begin{bmatrix} \boldsymbol{R}(0) & \boldsymbol{R}(1) & \cdots & \boldsymbol{R}(p-1) \\ \boldsymbol{R}(-1) & \boldsymbol{R}(0) & \cdots & \boldsymbol{R}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}(1-p) & \boldsymbol{R}(2-p) & \cdots & \boldsymbol{R}(0) \end{bmatrix}, \tag{6.90}$$

为以下递推的需要，现引入(6.89)的对偶方程

$$\boldsymbol{\Gamma}_p \mathscr{B}_p^{(p)} = \boldsymbol{\zeta}_p, \tag{6.91}$$

其中

$$\mathscr{B}_j^{(p)} = (\boldsymbol{B}_j^{(p)}, \boldsymbol{B}_{j-1}^{(p)}, \cdots, \boldsymbol{B}_1^{(p)})', \quad j = 1, 2, \cdots, p, \tag{6.92}$$

$$\boldsymbol{\zeta}_p = (\boldsymbol{R}'(p), \boldsymbol{R}'(p-1), \cdots, \boldsymbol{R}'(1))'. \tag{6.93}$$

当 $p=1$ 时，(6.89)变为

$$\boldsymbol{R}(0)(\boldsymbol{A}_1^{(1)})' = (\boldsymbol{R}(1))'.$$

相应地(6.91)变为

$$\boldsymbol{R}(0)(\boldsymbol{B}_1^{(1)})' = (\boldsymbol{R}'(1))'.$$

即 $p=1$ 时有

$$\boldsymbol{A}_1^{(1)} = \boldsymbol{R}(1)\boldsymbol{R}^{-1}(0), \qquad \boldsymbol{B}_1^{(1)} = \boldsymbol{R}'(1)\boldsymbol{R}^{-1}(0). \tag{6.94}$$

假设对第 p 步已求得 $\mathscr{A}_p^{(p)}$ 和 $\mathscr{B}_p^{(p)}$，则由(6.89)及(6.90)可得：

$$\begin{bmatrix} \boldsymbol{\Gamma}_p & \boldsymbol{\zeta}_p \\ \boldsymbol{\zeta}_p' & \boldsymbol{R}(0) \end{bmatrix} \begin{bmatrix} \mathscr{A}_p^{(p+1)} \\ (\boldsymbol{A}_{p+1}^{(p+1)})' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_p \\ \boldsymbol{R}'(p+1) \end{bmatrix}, \tag{6.95}$$

$$\begin{bmatrix} \boldsymbol{R}(0) & \boldsymbol{\eta}_p' \\ \boldsymbol{\eta}_p & \boldsymbol{\Gamma}_p \end{bmatrix} \begin{bmatrix} (\boldsymbol{B}_{p+1}^{(p+1)})' \\ \mathscr{B}_p^{(p+1)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}(p+1) \\ \boldsymbol{\zeta}_p \end{bmatrix}. \tag{6.96}$$

由(6.95)的第一组方程可得

$$\begin{aligned} &\boldsymbol{\Gamma}_p \mathscr{A}_p^{(p+1)} + \boldsymbol{\zeta}_p (\boldsymbol{A}_{p+1}^{(p+1)})' = \boldsymbol{\eta}_p, \\ &\mathscr{A}_p^{(p+1)} = \boldsymbol{\Gamma}_p^{-1}(\boldsymbol{\eta}_p - \boldsymbol{\zeta}_p (\boldsymbol{A}_{p+1}^{(p+1)})') \\ &\qquad = \mathscr{A}_p^{(p)} - \mathscr{B}_p^{(p)} (\boldsymbol{A}_{p+1}^{(p+1)})'. \end{aligned} \tag{6.97}$$

此即为

$$(\boldsymbol{A}_j^{(p+1)})' = (\boldsymbol{A}_j^{(p)})' - (\boldsymbol{B}_{p+1-j}^{(p)})'(\boldsymbol{A}_{p+1}^{(p+1)})', \tag{6.98}$$

$$j = 1, 2, \cdots, p.$$

将(6.97)代入(6.95)的第二组方程可得

$$\zeta'_p \mathscr{A}_p^{(p+1)} - \boldsymbol{R}(0)(\boldsymbol{A}_{p+1}^{(p+1)})' = \boldsymbol{R}'(p+1),$$

$$\zeta'_p(\mathscr{A}_p^{(p)} - \mathscr{B}_p^{(p)}(\boldsymbol{A}_{p+1}^{(p+1)})') + \boldsymbol{R}(0)(\boldsymbol{A}_{p+1}^{(p+1)})' = \boldsymbol{R}'(p+1),$$

$$\zeta'_p \mathscr{A}_p^{(p)} + (\boldsymbol{R}(0) - \zeta'_p \mathscr{B}_p^{(p)})(\boldsymbol{A}_{p+1}^{(p+1)})' = \boldsymbol{R}'(p+1).$$

可见

$$(\boldsymbol{A}_{p+1}^{(p+1)})' = (\boldsymbol{R}(0) - \zeta'_p \mathscr{B}_p^{(p)})^{-1}(\boldsymbol{R}'(p+1) - \zeta'_p \mathscr{A}_p^{(p)}). \tag{6.99}$$

具体写出来就是：

$$\boldsymbol{A}_{p+1}^{(p+1)} = \left(\boldsymbol{R}(p+1) - (\boldsymbol{A}_1^{(p)}, \cdots, \boldsymbol{A}_p^{(p)}) \begin{pmatrix} \boldsymbol{R}(p) \\ \vdots \\ \boldsymbol{R}(1) \end{pmatrix} \right)$$

$$\times \left(\boldsymbol{R}(0) - (\boldsymbol{B}_p^{(p)}, \cdots, \boldsymbol{B}_1^{(p)}) \begin{pmatrix} \boldsymbol{R}(p) \\ \vdots \\ \boldsymbol{R}(1) \end{pmatrix} \right)^{-1},$$

$$\begin{aligned} \boldsymbol{A}_{p+1}^{(p+1)} &= \left(\boldsymbol{R}(p+1) - \sum_{k=1}^{p} \boldsymbol{A}_k^{(p)} \boldsymbol{R}(p+1-k) \right) \\ &\quad \times \left(\boldsymbol{R}(0) - \sum_{j=1}^{p} \boldsymbol{B}_j^{(p)} \boldsymbol{R}(j) \right)^{-1}; \\ \boldsymbol{A}_j^{(p+1)} &= \boldsymbol{A}_j^{(p)} - \boldsymbol{A}_{p+1}^{(p+1)} \boldsymbol{B}_{p+1-j}^{(p)}, \qquad j = 1, 2, \cdots, p. \end{aligned} \tag{6.100}$$

又由(6.96)的第二组方程可得

$$\boldsymbol{\eta}_p (\boldsymbol{B}_{p+1}^{(p+1)})' + \boldsymbol{\Gamma}_p \mathscr{B}_p^{(p+1)} = \zeta_p, \tag{6.101}$$

或

$$\begin{aligned} \mathscr{B}_p^{(p+1)} &= \boldsymbol{\Gamma}_p^{-1}(\zeta_p - \boldsymbol{\eta}_p(\boldsymbol{B}_{p+1}^{(p+1)})') \\ &= \mathscr{B}_p^{(p)} - \mathscr{A}_p^{(p)}(\boldsymbol{B}_{p+1}^{(p+1)})'. \end{aligned} \tag{6.102}$$

因而

$$B_j^{(p+1)} = B_j^{(p)} - B_{p-1}^{(p+1)} A_{p+1-j}^{(p)}. \tag{6.103}$$

又由(6.96)的第一个方程得

$$R(0)(B_{p+1}^{(p+1)})' + \eta_p' \mathscr{B}_p^{(p+1)} = R(p+1),$$

$$R(0)(B_{p+1}^{(p+1)})' + \eta_p'(\mathscr{B}_p^{(p)} - \mathscr{A}_p^{(p)}(B_{p+1}^{(p+1)})') = R(p+1),$$

即

$$(B_{p+1}^{(p+1)})' = (R(0) - \eta_p' \mathscr{A}_p^{(p)})^{-1}(R(p+1) - \eta_p' \mathscr{B}_p^{(p)}),$$

$$B_{p+1}^{(p+1)} = \left(R'(p+1) - \sum_{j=1}^{p} B_j^{(p)} R'(p+j-1)\right)$$

$$\times\left(R(0) - \sum_{j=1}^{p} A_j^{(p)} R'(j)\right)^{-1}. \tag{6.104}$$

然而

$$\begin{aligned}
&(R'(p+1) - (\mathscr{B}_p^{(p)})'\eta_p) \\
&= (R'(p+1) - \zeta_p' \Gamma_p^{-1} \eta_p) \\
&= (R'(p+1) - \zeta_p' \mathscr{A}_p^{(p)}) \\
&= (R(p+1) - (\mathscr{A}_p^{(p)})' \zeta_p)' \\
&= \left(R(p+1) - \sum_{j=1}^{p} A_j^{(p)} R(p+1-j)\right)',
\end{aligned} \tag{6.105}$$

故(6.104)可用(6.105)改写为

$$B_{p+1}^{(p+1)} = \left(R(p+1) - \sum_{j=1}^{p} A_j^{(p)} R(p+1-j)\right)'$$

$$\times\left(R(0) - \sum_{j=1}^{p} A_j^{(p)} R'(j)\right)^{-1}. \tag{6.106}$$

将(6.100),(6.103)和(6.106)合并在一起，即组成了多维 Yule-Walker 方程的以下一组递推公式：

$$A_1^{(1)} = R(1)R^{-1}(0),\quad B_1^{(1)} = R'(1)R(0),$$

$$D_0^{(1)} = R_0 - A_1^{(1)} R_0 (A_1^{(1)})',$$

$$
\begin{cases}
A_{p+1}^{(p+1)} = \left(R(p+1) - \sum_{k=1}^{p} A_k^{(p)} R(p+1-k)\right) \\
\qquad \times \left(R(0) - \sum_{j=1}^{p} B_j^{(p)} R(j)\right)^{-1}, \\
A_j^{(p+1)} = A_j^{(p)} - A_{p+1}^{(p+1)} B_{p+1-j}^{(p)}, \qquad j = 1,2,\cdots,p; \\
B_{p+1}^{(p+1)} = \left(R(p+1) - \sum_{j=1}^{p} A_j^{(p)} R(p+1-j)\right)' \\
\qquad \times \left(R(0) - \sum_{j=1}^{p} A_j^{(p)} R'(j)\right)^{-1}; \\
B_j^{(p+1)} = B_j^{(p)} - B_{p+1}^{(p+1)} A_{p+1-j}^{(p)}, \qquad j = 1,2,\cdots,p. \\
D_0^{(p+1)} = (I - A_{p+1}^{(p+1)} B_{p+1}^{(p+1)}) D_0^{(p)}
\end{cases}
\tag{6.107}
$$

在实际计算中为了避免对(6.107)中有关的矩阵求逆运算，可以改为求以下线性方程组而直接得到 $A_{p+1}^{(p+1)}$ 和 $B_{p+1}^{(p+1)}$，这样可大大提高运算速度。

$$
\begin{cases}
\left(R(0) - \sum_{j=1}^{p} B_j^{(p)} R(j)\right)(A_{p+1}^{(p+1)})' \\
\qquad = \left(R(p+1) - \sum_{k=1}^{p} A_k^{(p)} R(p+1-k)\right)'; \\
\left(R(0) - \sum_{j=1}^{p} A_j^{(p)} R'(j)\right)(B_{p+1}^{(p+1)})' \\
\qquad = \left(R(p+1) - \sum_{j=1}^{p} A_j^{(p)} R(p+1-j)\right)'.
\end{cases}
\tag{6.108}
$$

在电算机上计算上述 Yule-Walker 方程的程序框图如图 10.1 所示。

如果是从一组样本

$$
X_t = (x_1(t), x_2(t), \cdots, x_n(t))', \quad t = 1,2,\cdots,N \tag{6.109}
$$

出发，则 $R(k)$ 可用以下样本协方差阵估计：

$$
\hat{r}_k = N^{-1} \sum_{l=1}^{N-k} X_{k+l} X'_l. \tag{6.110}
$$

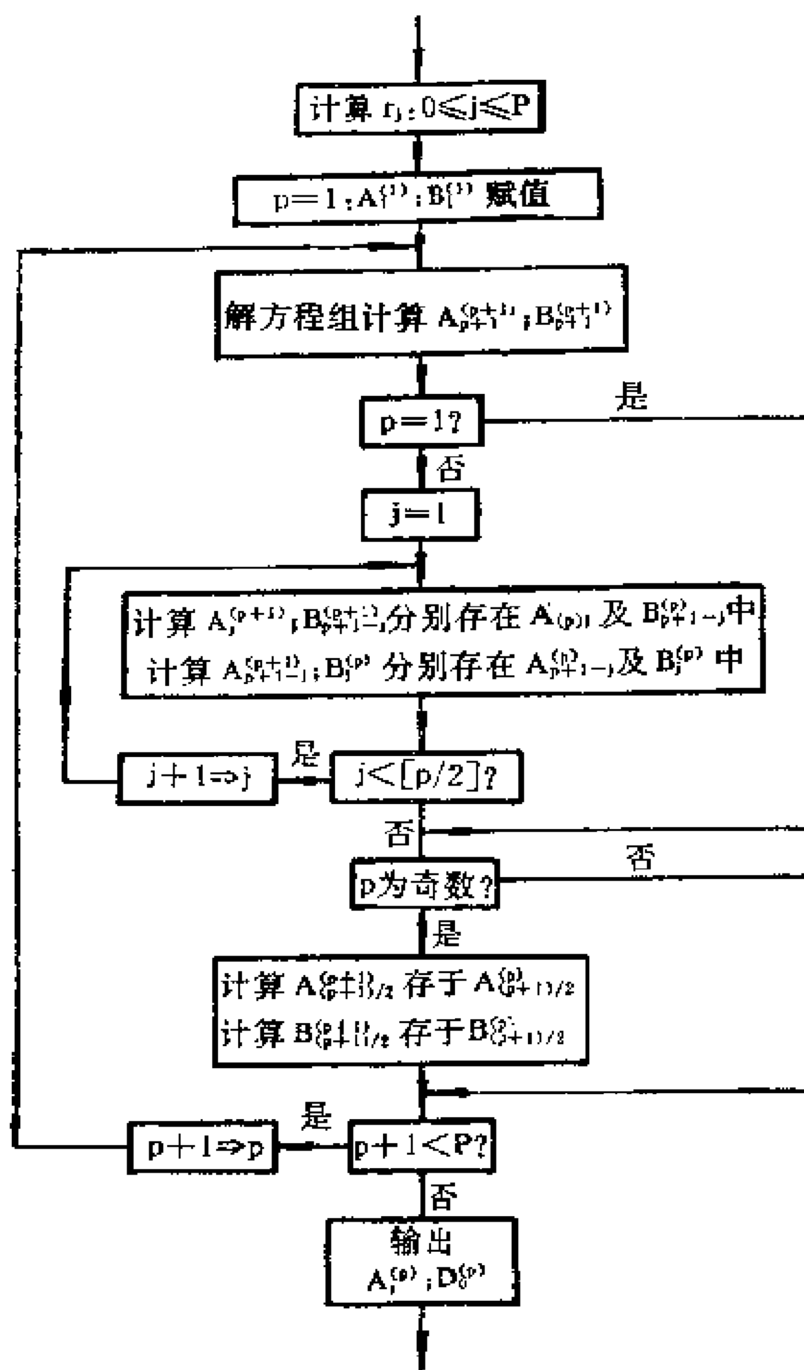


图 6.1

4° 多维 ARMA 模型与马氏扩张问题

读者知道在第二章的 § 4，我们曾详细地讨论了一维 ARMA 模型与马氏扩张问题。如今在 n 维条件下也有类似的结果。以下的叙述比第二章 § 4的说法要广一些，证明也更一般。

设

$$\boldsymbol{x}_t=(x_1(t),\cdots,x_m(t))', \quad t=0,\pm 1,\cdots$$

是实的宽平稳序列，记(必要时添加标记 x)：

$$H(t_0)=\mathscr{L}\{x_1(t_0),\cdots,x_n(t_0)\}, \quad t_0=0,\pm 1,\cdots, \tag{6.111}$$

$$H_{t^-}=\mathscr{L}\{x_i(s);\ s\leqslant t,\ i=1,2,\cdots,n\}, \tag{6.112}$$

$$H_{t^+}=\mathscr{L}\{x_i(s);\ s\geqslant t,\ i=1,2,\cdots,n\}, \tag{6.113}$$

$$H_t^{+|-}=\mathscr{L}\{\zeta;\ \zeta=\underset{H_{t^-}}{\mathrm{Proj}}(\xi),\ \xi\in H_{t^+}\}, \tag{6.114}$$

$$H_t^{-|+}=\mathscr{L}\{\zeta;\ \zeta=\underset{H_{t^+}}{\mathrm{Proj}}(\xi),\ \xi\in H_{t^-}\}. \tag{6.115}$$

定义6.4 称一个 n 维平稳列是马氏序列，若

$$H_t^{+|-}=H(t), \quad t=0,\pm 1,\cdots. \tag{6.116}$$

对给定的 n 维平稳列 $\boldsymbol{X}_t$ 如果存在一个 M 维平稳马氏列 $\boldsymbol{Y}_t=(y_1(t),\cdots,y_M(t))'$ 使得 $H^x(t_0)\subset H^y(t_0)$(指由(6.111)定义)则称 $\boldsymbol{Y}_t$ 是 $\boldsymbol{X}_t$ 的一个有限维马氏扩张；如果 $\boldsymbol{Y}_t$ 是 $\boldsymbol{X}_t$ 的一个有限维马氏扩张，并且对 $\boldsymbol{X}_t$ 的任一有限维马氏扩张 $\boldsymbol{Z}_t$，皆有

$$\dim H^y(t)\leqslant \dim H^z(t), \tag{6.117}$$

则称 $\boldsymbol{Y}_t$ 是 $\boldsymbol{X}_t$ 的一个最小马氏扩张。

定理6.9 设 $\boldsymbol{X}_t$ 是 n 维平稳列，若 $\dim H_{x,t}^{+|-}=M<+\infty$，则 $H_{x,t}^{+|-}$ 的任何一组基 $(y_1(t),\cdots,y_M(t))'$ 必构成 $\boldsymbol{X}_t$ 的一个最小马氏扩张。

证明 记 $\boldsymbol{Y}_t=(y_1(t),\cdots,y_M(t))'$，$\boldsymbol{Y}_t$ 是 $H_{x,t}^{+|-}$ 的一组基，

$$y_i(s)\in H^y(s)=H_{x,s}^{+|-}\subset H_{s^-}^x\Rightarrow H_{t^-}^y\subset H_{t^-}^x. \tag{6.118}$$

对 $s\leqslant t$

又因为

$$H^x(s)\subset H_{x,s}^{+|-}=H^y(s)\Rightarrow H_{t^-}^x\subset H_{t^-}^y,\ (s\leqslant t) \tag{6.119}$$

所以

$$H_{t^-}^y=H_{t^-}^x. \tag{6.120}$$

简记

$$P_A(B)=\{\underset{A}{\mathrm{Proj}}(\zeta);\ \zeta\in B\}, \tag{6.121}$$

其线性闭包为

$$[P_A(B)]=\mathscr{L}\{\zeta;\ \zeta\in P_A(B)\}, \tag{6.122}$$

则对任何 $u\geqslant t$ 皆有

$$P_{H^y_{t-}}(H^y(u))=P_{H^y_{t-}}(H^{+|-}_{x,u})=P_{H^y_{t-}}(P_{H^x_{u-}}(H^x_{u+})),$$

再由(6.120)，

$$\begin{aligned}P_{H^y_{t-}}(H^y(u)) &= P_{H^x_{t-}}(P_{H^x_{u-}}(H^x_{u+}))\ (\text{由 } H^y(u)=H^{+|-}_{x,u})\\ &= P_{H^x_{t-}}(H^x_{u+})\\ &\subset[P_{H^x_{t-}}(H^x_{u+})]\subset[P_{H^x_{t-}}(H^x_{t+})]=H^{+|-}_{x,t}\\ &=H^y(t).\end{aligned}\tag{6.123}$$

故

$$P_{H^y_{t-}}(H^y_{t+})\subset H^y(t),\quad H^{+|-}_{y,t}\subset H^y(t),\tag{6.124}$$

即

$$H^y(t)=H^{+|-}_{y,t}\quad(\text{因为 } H^y(t)\subset H^{+|-}_{y,t}).\tag{6.125}$$

表明 $\boldsymbol{Y}_t$ 是马氏列，并且是 $\boldsymbol{X}_t$ 的扩张，因为

$$H^x(t)\subset H^{+|-}_{x,t}=H^y(t).$$

以下证 $\boldsymbol{Y}_t$ 是最小马氏扩张。

设 $\boldsymbol{Z}_t$ 是 $\boldsymbol{X}_t$ 的另一个有限维马氏扩张，即

$$H^x(t)\subset H^z(t),\tag{6.126}$$

则

$$H^x_{t-}\subset H^z_{t-};\quad H^x_{t+}\subset H^z_{t+}.\tag{6.127}$$

所以

$$H^z(t)=H^{+|-}_{z,t}\supset P_{H^z_{t-}}(H^x_{t+}).$$

而

$$\begin{aligned}\dim P_{H^x_{t-}}(H^x_{t+}) &\leqslant \dim P_{H^z_{t-}}(H^x_{t+})\\ &\leqslant \dim H^z(t)=\dim H^{+|-}_{z,t},\end{aligned}\tag{6.127'}$$

从而

$$\dim H^y(t)\leqslant \dim H^z(t).$$

定理得证。 ▎

推论 $\boldsymbol{X}_t$ 具有有限维马氏扩张的充分必要条件是 $\dim H^{+|-}_t=M<+\infty$。

证明 若有有限维马氏扩张，则由(6.127)′

$$\dim H^{+|-}_{x,t} \leqslant \dim H^{+|-}_{x,t} = M < +\infty. \tag{6.128}$$

至于充分性可由定理6.9得出。▌

与一维的情形类似，运用第二章习题10—14的有关概念及定理6.9可得到有关马氏扩张与多维 ARMA 模型的关系。首先，一个 n 维平稳列 $\boldsymbol{X}_t$ 称为正则的，如果(见第二章习题13)

$$\bigcap_t H^{x-}_t = \varnothing. \tag{6.129}$$

定理6.10 对于正则的 n 维平稳列 $\boldsymbol{X}_t$，

$$\dim H^{+|-}_{x,t} = M < +\infty$$

的充要条件是 $\boldsymbol{X}_t$ 的谱密度阵 $\boldsymbol{f}(\lambda)$ 的每一元素 $f_{ij}(\lambda)$ 皆为 $e^{-i\lambda}$ 的有理函数。

定理6.11 设 $\boldsymbol{X}_t$ 是 n 维平稳列，它有有限维马氏扩张的充要条件是：$\boldsymbol{X}_t$ 的所有谱函数 $F_{ij}(\lambda)$ 都是 $e^{-i\lambda}$ 的有理函数的积分加上一个只在有限个点上有阶跃的函数。

读者不难看出定理6.11是第二章定理2.16的多维形式；而定理6.10(结合定理6.9)表明正则平稳列 $\boldsymbol{X}_t$ 可马氏扩张的充要条件乃是它为 ARMA 模型。

与第二章的问题相仿，有兴趣的问题是(定理6.9)给了具体的 n 维序列 $\boldsymbol{X}_t$(比如其谱密度阵)如何求出 $H^{+|-}_{x,t}$ 中的一组基 $(y_1(t), \cdots, y_m(t))'$？并且关心能否将这种最小的马氏扩张给予明显的表达(即能用谱密度阵 $\boldsymbol{f}(\lambda)$ 的参数来表示)。

以上问题以及定理6.10，6.11的证明，读者可参看[80]，[81]。

§3 多维时间序列的谱分析

在大量的现实数据分析中，多维平稳列的谱分析问题占有极其重要的地位，这方面的专著可参看[12]，[82]。以下仍分参数和非参数方法来介绍。

1° 多维极大熵准则下的模型拟合和谱估计

第四章中我们曾指出用极大一步预测误差的准则可代替极大熵准则并导致相同的结果。但前者可以免去事先假定考察的对象只限于 Shannon 熵存在或正则序列等的范围内。在多维的场合，运用极大一步预测误差准则就更有必要，因为就算谱密度阵 $\boldsymbol{f}(\lambda)$ 存在的条件下也可能是退化的——而在大量的实际问题中这种场合是不可迴避的。

考查 n 维实宽平稳列 $\boldsymbol{x}_t=(x_1(t),\cdots,x_n(t))'$，其相关阵为

$$\boldsymbol{R}_x(k)=E(\boldsymbol{x}_{t+k}\boldsymbol{x}'_t);\quad \boldsymbol{R}_x(-k)=\boldsymbol{R}'_x(k).$$

记

$$\boldsymbol{R}_x^{(M)}=\begin{bmatrix}\boldsymbol{R}_x(0) & \boldsymbol{R}_x(1) & \cdots & \boldsymbol{R}_x(M-1)\\ \boldsymbol{R}_x(-1) & \boldsymbol{R}_x(0) & \cdots & \boldsymbol{R}_x(M-2)\\ \cdots & \cdots & \cdots & \cdots\\ \boldsymbol{R}_x(1-M) & \boldsymbol{R}_x(2-M) & \cdots & \boldsymbol{R}_x(0)\end{bmatrix}_{(nM)\times(nM)}\tag{6.130}$$

$$=E\left\{\begin{pmatrix}\boldsymbol{x}(M)\\ \boldsymbol{x}(M-1)\\ \vdots\\ \boldsymbol{x}(0)\end{pmatrix}(\boldsymbol{x}'(M),\cdots,\boldsymbol{x}'(0))\right\},\tag{6.131}$$

其中第(j,k)元素为

$$(\boldsymbol{R}_x^{(M)})_{j,k}=\boldsymbol{R}_x(k-j).\tag{6.132}$$

令

$$H^x_{1,M}(t-1)=\mathscr{L}\{x_i(t-s)\,|\,i=1,2,\cdots,n,s=1,2,\cdots,M\},\tag{6.133}$$

$$\underset{H^x_{1,M}(t-1)}{\mathrm{Proj}}(\boldsymbol{x}(t))=\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{x}_{t-j}.\tag{6.134}$$

和(2.118)—(2.120)相仿，$\{\boldsymbol{B}_j^{(M)},1\leqslant j\leqslant M\}$应满足以下 Yule-Walker 方程：

$$\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\cdot\boldsymbol{R}_x(s-j)=\boldsymbol{R}_x(s).\tag{6.135}$$

用矩阵写法(参看6.89)):

$$\boldsymbol{R}_x^{(M)}\cdot\begin{pmatrix}(\boldsymbol{B}_1^{(M)})'\\(\boldsymbol{B}_2^{(M)})'\\\vdots\\(\boldsymbol{B}_M^{(M)})'\end{pmatrix}=\begin{pmatrix}\boldsymbol{R}_x'(1)\\\boldsymbol{R}_x'(2)\\\vdots\\\boldsymbol{R}_x'(M)\end{pmatrix}. \tag{6.136}$$

值得注意的是：即使在$\boldsymbol{R}_x^{(M)}$退化，$\{\boldsymbol{B}_j^{(M)}\}$诸系数不唯一，但(6.134)投影量不仅存在却还是唯一的。因此误差量为

$$\boldsymbol{D}_M=E\left\{\left(\boldsymbol{x}(t)-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{x}_{t-j}\right)\left(\boldsymbol{x}_t-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{x}_{t-j}\right)'\right\} \tag{6.137}$$

$$=\boldsymbol{R}_x(0)-(\boldsymbol{B}_1^{(M)},\cdots,\boldsymbol{B}_M^{(M)})\,\boldsymbol{R}^{(M)}\begin{pmatrix}(\boldsymbol{B}_1^{(M)})'\\\vdots\\(\boldsymbol{B}_M^{(M)})'\end{pmatrix}$$

$$=\boldsymbol{R}_x(0)-\sum_{k=1}^{M}\boldsymbol{B}_k^{(M)}\boldsymbol{R}_x(-k), \tag{6.138}$$

或

$$\boldsymbol{D}_M=\boldsymbol{D}_M'=\boldsymbol{R}_x(0)-\sum_{k=1}^{M}\boldsymbol{R}_x(k)(\boldsymbol{B}_k^{(M)})'. \tag{6.139}$$

如今问题的提法与一维相似：

设给了某个n维宽平稳列的头M个相关阵：

$$\{\boldsymbol{R}(k);\ k=0,1,2,\cdots,M\}\quad 且\ \boldsymbol{R}^{(M+1)}\geqslant 0, \tag{6.140}$$

要求一个拟合模型(也是n维的)$\boldsymbol{Y}_t$，使得：

(a) $\boldsymbol{R}_y(k)=\boldsymbol{R}(k),\quad k=0,1,2,\cdots,M.$ $\qquad$ (6.141)

(b) 在一切满足上述条件(a)之下，$\boldsymbol{Y}_t$的一步预测误差阵是最大的，即

$$\boldsymbol{D}_0=E\{\boldsymbol{Y}(0)-\underset{H_{-1}^{y}}{\mathrm{Proj}}(\boldsymbol{Y}(0))(\boldsymbol{Y}(0)-\underset{H_{-1}^{y}}{\mathrm{Proj}}(\boldsymbol{Y}(0)))'\} \tag{6.142}$$

是最大的，其中

$$H_{-1}^{y}=\mathscr{L}\{y_i(t);\ i=1,2,\cdots,n,t\leqslant -1\}. \tag{6.143}$$

和一维的场合相类似(见第二章习题10—14)，非奇异平稳列$\boldsymbol{Y}_t$可有如下展式(见[29])：

$$\boldsymbol{Y}_t=\sum_{k=0}^{\infty}\boldsymbol{C}_k\varepsilon_{t-k}+\boldsymbol{S}_t, \tag{6.144}$$

其中第一部分是 Wold 分解，$\boldsymbol{S}_t$ 为奇异部分，而 $\boldsymbol{\varepsilon}_t$ 是 $n_0\times 1$ 的白噪声(Wold)序列，$n_0\leqslant n$，$\boldsymbol{C}_k$ 是 $n\times n_0$ 的 Wold 系数阵。因而

$$\boldsymbol{D}_0=\boldsymbol{C}_0\boldsymbol{C}_0'. \tag{6.145}$$

记

$$\boldsymbol{P}(t)=\boldsymbol{Y}_t-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{Y}_{t-j} \tag{6.146}$$

$$=\boldsymbol{C}_0\boldsymbol{\varepsilon}_t+\sum_{k=1}^{\infty}\boldsymbol{d}_k\boldsymbol{\varepsilon}_{t-k}+\boldsymbol{\alpha}_t^{(s)}\quad(\text{利用}(6.144)).$$

于是

$$\begin{aligned}\boldsymbol{D}_M&=E\{\boldsymbol{P}(t)\boldsymbol{P}'(t)\}\\&=\boldsymbol{C}_0\boldsymbol{C}_0'+\sum\nolimits_{k=1}^{\infty}\boldsymbol{d}_k\boldsymbol{d}_k'+E\{\boldsymbol{\alpha}_t^{(s)}(\boldsymbol{\alpha}_t^{(s)})'\}.\end{aligned} \tag{6.147}$$

因而必得

$$\boldsymbol{D}_M\geqslant\boldsymbol{C}_0\boldsymbol{C}_0'. \tag{6.148}$$

而等号成立的充分必要条件显然应是

$$\boldsymbol{d}_k\equiv 0,\ \forall k\geqslant 1;\ \text{且}\ \boldsymbol{\alpha}_t^{(s)}=0,\ \forall t. \tag{6.149}$$

定理6.12(存在性) 设给了 $n\times n$ 的方阵 $\{\boldsymbol{R}(k)\}_0^M$，满足：

(1) $\boldsymbol{R}(k)=\boldsymbol{R}(-k)$，$k=0,1,\cdots,M$；

(2) $\boldsymbol{R}^{(M+1)}\geqslant 0$ (定义见(6.130))，

则存在唯一的一个 n 维的 AR(M)序列 $\boldsymbol{Y}_t$①，满足

$$\boldsymbol{R}_y(k)=\boldsymbol{R}(k),\quad k=0,1,\cdots,M.$$

其模型方程为

$$\boldsymbol{Y}_t-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{Y}_{t-j}=\boldsymbol{\varepsilon}_t, \tag{6.150}$$

其中 $\{\boldsymbol{B}_j^{(M)}\}_0^M$ 满足以下的 Yule-Walker 方程

$$\boldsymbol{R}_y^{(M)}\begin{pmatrix}(\boldsymbol{B}_1^{(M)})'\\ \vdots\\ (\boldsymbol{B}_M^{(M)})'\end{pmatrix}=\begin{pmatrix}\boldsymbol{R}_y(-1)\\ \vdots\\ \boldsymbol{R}_y(-M)\end{pmatrix}. \tag{6.151}$$

---

① 唯一性指 $\boldsymbol{Y}_t$ 的相关函数阵唯一。

而 ε_t 是 Y_t 的 Wold 序列，且

$$E\{\varepsilon_t\varepsilon_t'\}=D_M, \qquad (6.152)$$

其中

$$D_M=E\{P(t)P'(t)\}. \qquad (6.153)$$

$$P(t)=Y_t-\underset{H^y_{1,M}(t-1)}{\mathrm{Proj}}(Y(t)), \qquad (6.154)$$

$$H^y_{1,M}(t-1)=\mathscr{L}\{y_i(t-s);s=\overline{1,M},\quad i=\overline{1,n}\}. \quad (6.155)$$

证明 其方法类似于定理4.4，此处从略。 ▌

定理6.13（极大性） 设 $n\times n$ 的方阵 $R(k),k=0,1,\cdots,M$ 满足：

$$R(-k)=R'(k),\quad k=0,1,\cdots,M,\text{且 }R^{(M+1)}\geqslant 0.$$

考虑一切 n 维宽平稳列 Y_t，其相关函数矩阵满足

$$R_y(k)=R(k),\quad k=0,1,\cdots,M. \qquad (6.156)$$

为了使 Y_t 的一步预测误差成为最大，必须且只须 Y_t 是AR(M)模型。

证明 如果 Y_t 满足(6.156)并使 $C_0^{(y)}C_0^{(y)\prime}$ 达极大．则由定理6.12的存在性定理必有 AR(M)序列$\{Z_t\}$使

$$D_M^{(z)}=C_0^{(z)}C_0^{(z)\prime}.$$

在(6.156)条件下，对由(6.151)的一组系数矩阵$\{B_j^{(M)}\}$必有

$$\left\|Y_t-\sum_{j=1}^{M}B_j^{(M)}Y_{t-j}\right\|^2=D_M^{(y)}=D_M^{(z)}=C_0^{(z)}C_0^{(z)\prime}. \qquad (6.157)$$

由(6.146)，(6.148)知

$$C_0^{(z)}C_0^{(z)\prime}=D_M^{(y)}\geqslant C_0^{(y)}(C_0^{(y)})'.$$

由 Y_t 之极大性知

$$\left\|Y_t-\sum_{j=1}^{M}B_j^{(M)}Y_{t-j}\right\|^2=C_0^{(y)}C_0^{(y)\prime}.$$

由(6.149)，(6.146)知 Y_t 是 AR(M)。

反之，若 Y_t 是 AR(M)，则有

$$\boldsymbol{R}_y(M+l)-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{R}_y(M+l-j)=0,$$

且

$$\boldsymbol{C}_0^{(y)}\boldsymbol{C}_0^{(y)\prime}=\boldsymbol{D}_M^{(y)}. \tag{6.158}$$

对任何共平稳列 $\boldsymbol{Z}_t$，满足

$$\boldsymbol{R}_z(k)=\boldsymbol{R}(k),\quad k=0,1,\cdots,M,$$

必有

$$\left\|\boldsymbol{Z}_t-\sum_{j=1}^{M}\boldsymbol{B}_j^{(M)}\boldsymbol{Z}_{t-j}\right\|^2=\boldsymbol{D}_M^{(y)}\geqslant\boldsymbol{C}_0^{(z)}\cdot\boldsymbol{C}_0^{(z)\prime}. \tag{6.159}$$

表明 $\boldsymbol{C}_0^{(y)}\cdot\boldsymbol{C}_0^{(y)\prime}$ 极大。 |

根据以上所叙述的结果，对一个 n 维的(零均值)观测

$$\boldsymbol{X}_1,\ \boldsymbol{X}_2,\ \cdots,\ \boldsymbol{X}_N$$

可根据以下步骤来建模和谱估计：

(1) 先求样本协方差阵

$$\gamma_k=N^{-1}\sum_{j=1}^{N-k}\boldsymbol{X}_{j+k}\boldsymbol{X}'_j \tag{6.160}$$

且

$$\gamma'_k=\gamma_{-k}. \tag{6.161}$$

(2) 记

$$\boldsymbol{\Gamma}^{(M)}=\begin{bmatrix}\gamma_0 & \gamma_1 & \cdots & \gamma_{M-1}\\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{M-2}\\ \cdots & \cdots & \cdots & \cdots\\ \gamma_{1-M} & \gamma_{2-M} & \cdots & \gamma_0\end{bmatrix}. \tag{6.162}$$

用(6.107)及图6.1的流程解 Yule-Walker 方程

$$\begin{cases}\boldsymbol{\Gamma}^{(M)}\mathscr{A}_M^{(M)}=\eta_M,\\ \boldsymbol{\Gamma}^{(M)}\mathscr{B}_M^{(M)}=\zeta_M,\end{cases}\quad 1\leqslant M\leqslant L(N) \tag{6.163}$$

可得

$$(\boldsymbol{D}_0^{(M)};\boldsymbol{A}_1^{(M)},\cdots,\boldsymbol{A}_M^{(M)}),\quad 1\leqslant M\leqslant L(N),$$

其中 $L(N)$是与N有关的上界，可选 $O(\ln N)$或实用的 $\sqrt{N}$等。

(3) 用 AIC 来定阶：即选使得

$$\mathrm{AIC}(M)=\ln\det(\boldsymbol{D}_0^{(M)})+\frac{2}{N}(Mn^2),$$

其中
$$1\leqslant M\leqslant L(N) \tag{6.164}$$
达极小的M值，记为 p_0 作为拟合的阶。

可以证明：当 $\boldsymbol{\Gamma}^{(M+1)}>0$ 时，由

$$\boldsymbol{\Gamma}^{(M)}\mathscr{A}_M^{(M)}=\boldsymbol{\eta}_M$$

确定的矩阵系数$\{\boldsymbol{A}_1^{(M)},\cdots,\boldsymbol{A}_M^{(M)}\}$必使得

$$\det\left\{z^M\boldsymbol{I}_n-\sum_{j=1}^{M}\boldsymbol{A}_j^{(M)}z^{M-j}\right\}=0 \tag{6.165}$$

的根皆在单位圆内，或说

$$\det\left\{\sum_{j=0}^{M}\boldsymbol{A}_jz^j\right\}$$

的根皆在单位圆外，其中

$$\boldsymbol{A}_j=-\boldsymbol{A}_j^{(M)},\qquad \boldsymbol{A}_0=\boldsymbol{I}_n.$$

(4) 拟合的 n 维 AR(p_0)模型为

$$\boldsymbol{X}_t+\boldsymbol{A}_1\boldsymbol{X}_{t-1}+\cdots+\boldsymbol{A}_{p_0}\boldsymbol{X}_{t-p_0}=\boldsymbol{\varepsilon}_t,\quad t=0,\pm1,\cdots, \tag{6.166}$$

其中 $\boldsymbol{\varepsilon}_t$ 为白噪声列，$E\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'=\boldsymbol{D}_0^{(p_0)}$，并且谱阵的估计为

$$\hat{\boldsymbol{f}}(\lambda)=\frac{1}{2\pi}\boldsymbol{A}^{-1}(\mathrm{e}^{-i\lambda})\boldsymbol{D}_0^{(p_0)}\boldsymbol{A}(\mathrm{e}^{-i\lambda}),\qquad -\pi\leqslant\lambda\leqslant\pi, \tag{6.167}$$

其中
$$\boldsymbol{A}(\mathrm{e}^{-i\lambda})=\sum_{k=0}^{p_0}\boldsymbol{A}_k\mathrm{e}^{-ik\lambda}.$$

2° 多维加窗谱估计方法

设

$$\boldsymbol{X}_t=(x_1(t),x_2(t),\cdots,x_n(t))',\quad t=1,2,\cdots,N$$

为观测值($Ex_i(t)\equiv0$)，以下介绍分段平均求谱阵的方法。

设将N个资料分为互不重叠的等长的L段，第μ段的第j个分量的加窗 Fourier 变换为

$$I_{j,\mu}(\lambda) = \frac{1}{\sqrt{L}}\sum_{m=1}^{L} w_m x_j(m+(\mu-1)L)e^{-im\lambda}, \quad (6.168)$$

$$j=1,2,\cdots,n; \quad \mu=1,2,\cdots,M,$$

其中$\{w_m\}$为窗系数。为统一起见，可简单取

$$w_m \equiv 1, \quad m=1,2,\cdots,L, \quad (6.169)$$

于是可用

$$S_{jk}(\lambda) = \frac{1}{M}\sum_{\mu=1}^{M} I_{j\mu}(\lambda)\bar{I}_{k\mu}(\lambda), \quad (6.170)$$

$$j,k=1,2,\cdots,n$$

来估计$\boldsymbol{f}_x(\lambda)$中的$f_{jk}(\lambda)$。此时，

$$\begin{aligned}
E[S_{jk}(\lambda)] &= E\left[\frac{1}{M}\sum_{\mu=1}^{M} I_{j\mu}(\lambda)\bar{I}_{k\mu}(\lambda)\right] \\
&= E\left[\frac{1}{ML}\sum_{\mu=1}^{M}\sum_{m,m'=1}^{L} x_j(m+(\mu-1)L)\right. \\
&\quad \left.\times x_k(m'+(\mu-1)L)e^{-i(m-m')\lambda}\right] \\
&= \frac{1}{L}\sum_{m,m'=1}^{L} R_{jk}(m-m')e^{-i(m-m')\lambda} \\
&= \int_{-\pi}^{\pi} f_{jk}(u)\,|W_N(u-\lambda)|^2\frac{du}{2\pi},
\end{aligned} \quad (6.171)$$

其中$|W_N(\lambda)|^2$是 Bartlett 窗对应的 Fejer 函数：

$$|W_N(\lambda)|^2 = \frac{1}{N}\left|\frac{\sin\frac{N\lambda}{2}}{\sin\frac{\lambda}{2}}\right|^2, \quad (6.172)$$

并且

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|W_N(\lambda)|^2 d\lambda = 1.$$

利用第五章§3的结果和方法不难看出在不太严的条件下有

$$E[S_{jk}(\lambda)]\longrightarrow f_{jk}(\lambda)\quad (N\to\infty). \tag{6.173}$$

当 $\boldsymbol{X}_t$ 是 n 维正态的条件下，

$$\begin{aligned}
&E[x_j(m+(\mu-1)L)x_k(m'+(\mu-1)L)\\
&\quad\times x_j(m''+(\mu'-1)L)x_k(m'''+(\mu'-1)L)]\\
&= R_{jk}(m-m')R_{jk}(m''-m''')\\
&\quad+R_{jj}(m-m''+(\mu-\mu')L)\\
&\quad\times R_{kk}(m'-m'''+(\mu-\mu')L)\\
&\quad+R_{jk}(m-m'''+(\mu-\mu')L)\\
&\quad\times R_{kj}(m'-m''+(\mu-\mu')L),
\end{aligned}$$

从而

$$\begin{aligned}
&E[S_{jk}^2(\lambda_0)]\\
&\quad= E\left[\frac{1}{M^2L^2}\sum_{\mu,\mu'=1}^{M}\left(\sum_{m,m',m'',m'''=1}^{L}x_j(m+(\mu-1)L)\right.\right.\\
&\quad\times x_k(m'+(\mu-1)L)\cdot x_j(m''+(\mu'-1)L)\\
&\quad\left.\left.\times x_k(m'''+(\mu'-1)L)\cdot \mathrm{e}^{-\mathrm{i}(m-m'+m''-m''')\lambda}\right)\right].
\end{aligned}\tag{6.174}$$

因此方差

$$\begin{aligned}
D(S_{jk}(\lambda_0)) &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left\{\frac{1}{M}+\frac{2}{M}\sum_{\mu=1}^{M-1}\left(\frac{M-\mu}{M}\right)\mathrm{e}^{-\mathrm{i}\mu L(\lambda-\lambda')}[f_{jj}(\lambda)\right.\\
&\quad\times f_{kk}(\lambda')\,|W_L(\lambda-\lambda_0)W_L(\lambda'-\lambda_0)|^2\\
&\quad+f_{jk}(\lambda)f_{jk}(\lambda')W_L(\lambda-\lambda_0)\\
&\quad\times W_L^*(\lambda+\lambda_0)W_L(\lambda'+\lambda_0)\\
&\quad\left.\times W_L^*(\lambda-\lambda_0)]\right\}\frac{\mathrm{d}\lambda}{2\pi}\;\frac{\mathrm{d}\lambda'}{2\pi}.
\end{aligned}\tag{6.175}$$

当N充分大时，有以下渐近式成立：

$$D(S_{jk}(\lambda_0))\sim\begin{cases}\dfrac{1}{M}f_{jj}(\lambda_0)f_{kk}(\lambda_0), & \lambda_0\neq 0,\pm\pi;\\[2ex] \dfrac{1}{M}[f_{jj}(\lambda_0)f_{kk}(\lambda_0)+f_{jk}^2(\lambda_0)], & \text{否则}.\end{cases}\tag{6.176}$$

表明当$M\to\infty$时，$S_{jk}(\lambda)$是$f_{jk}(\lambda)$的相容估计。

以上介绍的是把窗加在观测数据上，这样做的好处是可直接利用快速 Fourier 变换。此外，亦可考虑把窗系数加到相关函数阵上。例如：

$$\hat{\boldsymbol{f}}_N(\lambda)=\frac{1}{2\pi}\sum_{s=1-N}^{N-1}w_N(s)\boldsymbol{\gamma}_s e^{-i\lambda s}, \tag{6.177}$$

其中$\boldsymbol{\gamma}_s$的定义如(6.160)，(6.161)，而$w_N(s)$是谱窗系数。当然(6.177)中也可以根据不同的分量选不同的窗，如

$$f_{k,j}^{(N)}(\lambda)=\frac{1}{2\pi}\sum_{s=1-N}^{N-1}w_{k,j}^{(N)}(s)\gamma_{k,j}(s)e^{-is\lambda}, \tag{6.177$'$}$$

其中

$$\begin{cases}\gamma_{k,j}(s)=N^{-1}\sum_{l=1}^{N-s}x_k(l+s)x_j(l), & s=0,1,\cdots,N-1;\\ \gamma_{k,j}(s)=N^{-1}\sum_{l=1-s}^{N}x_k(l+s)x_j(l), & s=-1,-2,\cdots,-N+1.\end{cases} \tag{6.178}$$

然而(6.177)$'$的加窗法就会大大增加计算量，因此目前大多数人仍选取统一的(6.177)的加窗方式。

§4 多维时间序列的预测

1° 有理谱阵预测问题的 Yaglom 方法

多维时间序列的预测和滤波问题在应用中是十分重要的。然而在理论上其解决的圆满程度还没有象一维的那样彻底。以下介绍比较实用的有理谱密度条件(即多维 ARMA 模型)下的预测方法，此方法也是属于 A.M.Yaglom 的(参看[4])。在[80]中得到比较清楚的整理，以下以$n=2$来说明。

首先指出以下讨论的条件是：

(a) $\boldsymbol{X}_t$平稳列的谱密度阵存在并且是$e^{i\lambda}$的有理函数，记为$\boldsymbol{f}(e^{i\lambda})$。

(b) $\det(\boldsymbol{f}(e^{i\lambda}))>0$, $-\pi\leqslant\lambda\leqslant\pi$, 其中

$$\boldsymbol{f}(e^{i\lambda})=\begin{pmatrix} f_{11}(e^{i\lambda}) & f_{12}(e^{i\lambda}) \\ f_{21}(e^{i\lambda}) & f_{22}(e^{i\lambda}) \end{pmatrix}. \tag{6.179}$$

(c) 记

$$\begin{pmatrix} f_{11}(e^{i\lambda}) & f_{12}(e^{i\lambda}) \\ f_{21}(e^{i\lambda}) & f_{22}(e^{i\lambda}) \end{pmatrix}^{-1}=\begin{pmatrix} f^*_{11}(e^{i\lambda}) & f^*_{12}(e^{i\lambda}) \\ f^*_{21}(e^{i\lambda}) & f^*_{22}(e^{i\lambda}) \end{pmatrix}.$$

定理6.14 设 $\boldsymbol{X}_t=(x_1(t),x_2(t))$是(零均值)平稳列，满足(a),(b)条件，则对 $\tau\geqslant 0$

$$\hat{x}_1(t+\tau)=\underset{H_x(t-1)}{\operatorname{Proj}}(x_1(t+\tau)) \tag{6.180}$$

$$=\int_\Pi e^{it\lambda}\Phi_1^-(e^{i\lambda})dz_1(\lambda)+\int_\Pi e^{it\lambda}\Phi_2^-(e^{i\lambda})dz_2(\lambda)$$

$$=\int_\Pi e^{it\lambda}(\Phi_1^-,\Phi_2^-)d\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \tag{6.181}$$

其中

$$H_x(t-1)=\mathscr{L}\{x_i(s);\ s\leqslant t-1, i=1,2\}, \tag{6.182}$$

而 dz_1,dz_2 分别为 $x_1(t),x_2(t)$的正交测度；Φ_1^-和 Φ_2^- 分别满足以下条件：

(i)
$$\int_\Pi|\Phi_1^-(e^{i\lambda})|^2f_{11}(e^{i\lambda})d\lambda<+\infty,$$
$$\int_\Pi|\Phi_2^-(e^{i\lambda})|^2f_{22}(e^{i\lambda})d\lambda<+\infty. \tag{6.183}$$

(ii) Φ_1^-,Φ_2^-都是 $e^{i\lambda}$ 的有理函数，且可写成

$$\begin{cases} \Phi_1^-(e^{i\lambda})=\dfrac{W_1^-(e^{i\lambda})}{(e^{i\lambda}-\theta_{11})\cdots(e^{i\lambda}-\theta_{1L_1})}, \\ \Phi_2^-(e^{i\lambda})=\dfrac{W_2^-(e^{i\lambda})}{(e^{i\lambda}-\theta_{21})\cdots(e^{i\lambda}-\theta_{2L_2})}, \end{cases} \tag{6.184}$$

其中$(\theta_{11},\cdots\theta_{1L_1})$是 $f^*_{11}(z),f^*_{21}(z)$的分母的最小公倍多项式在单位圆内全部的根；$(\theta_{21}\cdots\theta_{2L_2})$是 $f^*_{12}(z),f^*_{22}(z)$的分母的最小公倍多项式在单位圆内的全部根；$W_1^-(e^{i\lambda})$和 $W_2^-(e^{i\lambda})$分别是 $e^{i\lambda}$

的 L_1-1 次和 L_2-1 次多项式，$\Phi_1^-(z)$ 和 $\Phi_2^-(z)$ 在 $|z|\geqslant 1$ 上是解析的并且 $\Phi_1^-(\infty)=\Phi_2^-(\infty)=0$。

(iii) 适当选择 W_1^- 和 W_2^- 使得

$$\begin{cases}\Psi_1^-(z)=[z^\tau-\Phi_1^-(z)]f_{11}(z)-\Phi_2^-(z)f_{21}(z),\\ \Psi_2\ (z)=[z^\tau-\Phi_1^-(z)]f_{12}(z)-\Phi_2^-(z)f_{22}(z),\end{cases}\tag{6.185}$$

它们的分母的根都在单位圆外。

证明 首先我们指出：在条件(a)，(b)下，不难看出必有

$$\det(\boldsymbol{f}(e^{i\lambda}))\geqslant\delta>0,\quad \lambda\in\Pi,\tag{6.186}$$

并且

$$0<\alpha\leqslant f_{jj}(e^{i\lambda})\leqslant\beta,\quad j=1,2,\quad \lambda\in\Pi.\tag{6.187}$$

从而

$$\frac{\det(\boldsymbol{f}(e^{i\lambda}))}{\prod_{j=1}^{2}f_{jj}(e^{i\lambda})}\geqslant\frac{\delta}{\beta^2}>0.\tag{6.188}$$

由定理 6.5 知 $L^2(d\boldsymbol{F})$ 是可分离的，即 $\underset{\smile}{\varphi}=(\varphi_1,\varphi_2)\in L^2(d\boldsymbol{F})$ 必有 $\varphi_j\in L^2(dF_{jj})$，$j=1,2$。因此，对任意的 $\xi\in H_x(t-1)$，

$$\xi=\sum_{j=1}^{2}\int_\Pi\varphi_j^{(\xi)}(\lambda)dz_j(\lambda),\tag{6.189}$$

$\varphi_j^{(\xi)}(\lambda)\in L^2_{-1}(dF_{jj})$，$j=1,2$，即 $\varphi_j^{(\xi)}(\lambda)\in\mathscr{L}\{e^{-is\lambda};\ s\geqslant 1\}$。

由(6.187)知此时必有：若 $\varphi_j^{(N)}(\lambda)\xrightarrow{L^2(dF_{jj})}\varphi_j^{(\xi)}(\lambda)$，则

$$\int_\Pi|\varphi_j^{(N)}(\lambda)-\varphi_j^{(\xi)}(\lambda)|^2d\lambda\to 0\ \ (N\to\infty).$$

因此(在 L^2 意义下)

$$\varphi_j^{(\xi)}(\lambda)=\sum_{k=1}^{\infty}a_k^{(j)}e^{-ik\lambda},\quad j=1,2.\tag{6.190}$$

记

$$\hat{x}_1(t+\tau)=\underset{H_x(t-1)}{\mathrm{Proj}}(x_1(t+\tau))$$

$$= \int_{\Pi} e^{il\lambda} \Phi_1^-(e^{i\lambda}) dz_1 + \int_{\Pi} e^{il\lambda} \Phi_2^-(e^{i\lambda}) dz_2. \tag{6.191}$$

因为

$$x_1(t+\tau) - \hat{x}_1(t+\tau) \perp x_j(t-l), \quad j=1,2, \quad l=1,2,\cdots, \tag{6.192}$$

故有

$$\begin{cases} \int_{\Pi} \{[e^{i\tau\lambda} - \Phi_1^-(e^{i\lambda})] f_{11}(e^{i\lambda}) - \Phi_2^-(e^{i\lambda}) f_{21}(e^{i\lambda})\} e^{il\lambda} d\lambda = 0; \\ \int_{\Pi} \{[e^{i\tau\lambda} - \Phi_1^-(e^{i\lambda})] f_{12}(e^{i\lambda}) - \Phi_2^-(e^{i\lambda}) f_{22}(e^{i\lambda})\} e^{il\lambda} d\lambda = 0. \end{cases} \tag{6.193}$$

对 $l=1,2,\cdots$ 成立，其中

$$\Phi_j^-(e^{i\lambda}) = \sum_{k=1}^{\infty} a_k^{(j)} e^{-ik\lambda}, \quad j=1,2. \tag{6.194}$$

反之也对：即若 $\{\Phi_j^-(e^{i\lambda}), j=1,2\}$ 具(6.194)形式，并使(6.193)对 $l \geqslant 1$ 成立，$\Phi_j^-(e^{i\lambda}) \in L^2(dF_{jj})$，则由(6.191)表示的 $\hat{x}_1(t+\tau)$ 必是最优预报。

可见，对(6.184)只需验以上条件即可。首先由(6.184)分母的选法知 $\Phi_j^-(z)$ 在 $|z| \geqslant \rho(\rho<1)$ 上是解析的，$\Phi_j^-(\infty)=0, j=1, 2$，表明它们皆具有(6.194)的形式。又因 $|\Phi_j(z)|$ 在单位圆周上无根，所以都是有正上下界的函数，从而

$$\Phi_j^-(e^{i\lambda}) \in L^2(dF_{jj}), \quad j=1,2.$$

至于(6.193) 只要适当选择 $W_1^-(e^{i\lambda})$ 和 $W_2^-(e^{i\lambda})$ 必可使 $\Psi_1^-(z)$ 和 $\Psi_2(z)$ (见(6.185))分母的根都在单位圆外。显见 $\Psi_j^-(e^{i\lambda}) \in L^2(d\lambda)$ 并可表为

$$\Psi_j^-(e^{i\lambda}) = \sum_{0}^{\infty} c_k^{(j)} e^{ik\lambda}, \quad j=1,2.$$

于是(6.192)必满足，由此证明了(6.181)是最优预报。 ▍

2° 一般预测问题

一般的 n 维平稳列的预测理论已超出本书范围，可参看[29]，以下只介绍几个结果。

首先我们给出正则和奇异的概念。

定义6.4 设 $\boldsymbol{X}_t, t=0,\pm1,\cdots$ 为 n 维平稳序列，记

$$H_x(t)=\mathscr{L}\{x_i(s);\ i=\overline{1,n},\ s\leqslant t\}.$$

若

$$\bigcap_t H_x(t)=\varnothing, \tag{6.195}$$

则称 $\boldsymbol{X}_t$ 是正则的；若

$$\bigcap_t H_x(t)=H_x=\mathscr{L}\{x_i(t);\ i=\overline{1,n}, t=0,\pm1,\cdots\}, \tag{6.196}$$

则称 $\boldsymbol{X}_t$ 是奇异的。

设 $\boldsymbol{X}_t$ 是 n 维正则平稳列，用 $D_x(t)$ 表 $H_x(t)$ 中 $H_x(t-1)$ 的正交补：

$$D_x(t)=H_x(t)\ominus H_x(t-1), \tag{6.197}$$

则

$$H_x(t)=\sum_{s=-\infty}^{t}\oplus D_x(s), \tag{6.198}$$

并且

$$H_x=\sum_{s=-\infty}^{\infty}\oplus D_x(s). \tag{6.199}$$

这些结果在第二章中已在一维场合下介绍过。对 $D_x(0)$，设它的基为 ζ_j，$j=1,2,\cdots,m$，由(6.60)

$$\zeta_j=\int_\Pi \underset{\sim}{a}_j(\lambda)\mathrm{d}\boldsymbol{Z}_X(\lambda),\qquad j=\overline{1,m}.$$

记

$$\zeta_j(t)=\int_\Pi \mathrm{e}^{it\lambda}\,\underset{\sim}{a}_j(\lambda)\mathrm{d}\boldsymbol{Z}_X(\lambda),\qquad j=\overline{1,m}. \tag{6.200}$$

显见

$$\{\zeta_j(t);\ j=\overline{1,m},\ \ t=0,\pm 1,\cdots\} \tag{6.201}$$

组成了 H_x 中的标准正交基(称为基本列),从而有以下的 Wold分解式

$$\boldsymbol{X}_t=\sum_{s=-\infty}^{t}\boldsymbol{C}_{t-s}\boldsymbol{\zeta}(s), \tag{6.202}$$

其中

$$\boldsymbol{C}_t=(c_{k,j}(t))_{j=\overline{1,m};k=\overline{1,n}} \tag{6.203}$$

$$c_{kj}(t)=E(x_k(t)\overline{\zeta_j(0)}),\qquad k=\overline{1,n},\ j=\overline{1,m}, \tag{6.204}$$

$$x_k(t)=\sum_{s=-\infty}^{t}\sum_{j=1}^{m}c_{k,j}(t-s)\zeta_j(s),\qquad k=\overline{1,n}. \tag{6.205}$$

从而已知$\{\boldsymbol{X}_{t'};t'\leqslant t\}$对 $x_k(t+\tau)$(比如$k=1$), $\tau>0$ 的最优预测显然应是:

$$\underset{H_x(t)}{\mathrm{Proj}}\{x_k(t+\tau)\}. \tag{6.206}$$

利用$\{\boldsymbol{\zeta}(t)\}$的等时从属性

$$\zeta_j(t)\in H_x(t)\ominus H_x(t-1),\qquad j=\overline{1,m}. \tag{6.207}$$

由(6.214) 知

$$\hat{x}_k(t+\tau)=\sum_{s=-\infty}^{t}\sum_{j=1}^{m}c_{kj}(t+\tau-s)\zeta_j(s),\qquad k=\overline{1,n}, \tag{6.208}$$

并且误差

$$\begin{aligned}\Delta_{k,\tau}^2&=\|x_k(t+\tau)-\hat{x}_k(t+\tau)\|^2\\&=\sum_{s=t+1}^{t+\tau}\sum_{j=1}^{m}|c_{k,j}(t+\tau-s)|^2,\qquad k=\overline{1,n}.\end{aligned} \tag{6.209}$$

值得注意的是:和一维场合之下不同,基本列$\{\boldsymbol{\zeta}(t)\}$只能在差一个酉阵的意义下被确定,即设$\{\boldsymbol{\zeta}(t)\}$是 H_x 中的基本列,而 $\boldsymbol{U}$是 $m\times m$ 的任一酉阵,则

$$\tilde{\boldsymbol{\zeta}}(t)=\boldsymbol{U}\boldsymbol{\zeta}(t),\quad t=0,\pm 1,\cdots \tag{6.210}$$

仍然是 H_x 的基本列。因为

(a) $\tilde{\zeta}(t) \in H_x(t) \ominus H_x(t-1)$; (6.211)

(b) $E(\widetilde{\zeta}(t)\widetilde{\zeta}^*(s)) = \boldsymbol{U}E\zeta(t)\zeta^*(s)\boldsymbol{U}^*$ (6.212)

$$= \delta_{t,s}\boldsymbol{I}_{m\times m}. \tag{6.213}$$

因而 $\tilde{\zeta}(t)$ 仍为基本列。

可见问题的关键是如何能得到基本列 $\{\zeta(t)\}$ 并得到 Wold 系数 $\{c_{kj}(t)\}$。

我们称一个 n 维随机平稳列 $\boldsymbol{X}_t$ 是秩为 m 的 $(1\leqslant m\leqslant n)$，如果它的谱密度阵

$$\boldsymbol{f}(\lambda) = (f_{kj}(\lambda))_{j=\overline{1,n};\ k=\overline{1,n}} \tag{6.214}$$

对几乎所有的 $\lambda\in\Pi$ 都是秩为 m 的。

称一个解析函数阵

$$\boldsymbol{\Gamma}(z) = (\Gamma_{kj}(z))_{j=\overline{1,m};\ k=\overline{1,n}} \tag{6.215}$$

是 H_2 解析函数阵，如果：

它的每个元素 $\Gamma_{kj}(z)$ 皆是单位圆内的解析函数，并满足

$$\sup_{0\leqslant\rho<1}\int_{\Pi}|\Gamma_{kj}(\rho e^{-i\lambda})|^2 d\lambda < +\infty; \tag{6.216}$$

设 $\boldsymbol{X}_t$ 是 n 维的正则平稳列，它的秩为 m，称 $\boldsymbol{\Gamma}(z) = (\Gamma_{kj}(z))_{j=\overline{1,m};\ k=\overline{1,n}}$ 是单位圆内的 H_2 极大解析函数阵，若它使 $\boldsymbol{X}_t$ 的谱密度阵 $\boldsymbol{f}(\lambda)$ 可分解为

$$\boldsymbol{f}(\lambda) = \frac{1}{2\pi}\boldsymbol{\Gamma}(e^{-i\lambda})\boldsymbol{\Gamma}^*(e^{-i\lambda}), \tag{6.217}$$

并对任何满足上述条件的解析函数阵 $\widetilde{\boldsymbol{\Gamma}}(z)$，皆有

$$\boldsymbol{\Gamma}(0)\boldsymbol{\Gamma}^*(0) \geqslant \widetilde{\boldsymbol{\Gamma}}(0)\widetilde{\boldsymbol{\Gamma}}^*(0) \tag{6.218}$$

成立。不难看出，对极大阵 $\boldsymbol{\Gamma}(z)$ 必存在矩阵

$$\boldsymbol{\psi}(\lambda) = (\psi_{jk}(\lambda))_{k=\overline{1,n};\ j=1,m}$$

使得

$$\psi(\lambda)\boldsymbol{\Gamma}(e^{-i\lambda})=\boldsymbol{I}_m. \tag{6.219}$$

于是

$$\zeta(t)=\int_\Pi e^{i\lambda t}\psi(\lambda)d\boldsymbol{Z}_x(\lambda),\quad t=0,\pm 1,\cdots \tag{6.220}$$

必是基本列，因为

$$\begin{aligned}E\zeta(t)\zeta^*(s)&=\int_\Pi e^{i(t-s)\lambda}\psi(\lambda)\boldsymbol{f}(\lambda)\psi^*(\lambda)d\lambda\\&=\frac{1}{2\pi}\int_\Pi e^{i(t-s)\lambda}(\psi\boldsymbol{\Gamma}(e^{-i\lambda}))(\psi\boldsymbol{\Gamma}(e^{-i\lambda}))^*d\lambda\\&=\delta_{t,s}\boldsymbol{I}_{m\times m}.\end{aligned} \tag{6.221}$$

而

$$\Gamma_{k,j}(z)=\sum_{t=0}^{\infty}c_{k,j}(t)z^t \tag{6.222}$$

的系数$\{c_{k,j}(t)\}$即为(6.213)的系数。

于是问题转为给了$\boldsymbol{X}_t$的谱密度阵$\boldsymbol{f}(\lambda)$之后如何去找极大解析函数阵$\boldsymbol{\Gamma}(z)$。当$\boldsymbol{f}(\lambda)$是有理谱密度阵，或$\boldsymbol{f}(\lambda)$满足

$$c_1\boldsymbol{I}_{n\times n}\leqslant\boldsymbol{f}(\lambda)\leqslant c_2\boldsymbol{I}_{n\times n}, \tag{6.223}$$

其中$0<c_1\leqslant c_2<+\infty$。[29]中给出了如何寻找极大$\boldsymbol{\Gamma}(z)$的想法和步骤，但都相当繁杂，很难实用，目前可以说还没有圆满的解法。

最后，如果$\boldsymbol{X}_t$是正则的，秩也是n，则称为是满秩正则的。如果$\boldsymbol{\Gamma}(z)$是H_2解析函数阵(单位圆内)，满足

$$\boldsymbol{f}(\lambda)=\frac{1}{2\pi}\boldsymbol{\Gamma}(e^{-i\lambda})\boldsymbol{\Gamma}^*(e^{-i\lambda})$$

则$\boldsymbol{\Gamma}(z)$是极大的充分必要条件是

$$|\det\boldsymbol{\Gamma}(0)|^2=(2\pi)^n\exp\left\{\frac{1}{2\pi}\int_\Pi\log\det\boldsymbol{f}(\lambda)d\lambda\right\}. \tag{6.224}$$

而$\boldsymbol{X}_t$正则的条件下，它是满秩的，当且仅当

$$\int_\Pi\log\det\boldsymbol{f}(\lambda)d\lambda>-\infty. \tag{6.225}$$

参考文献和书目

[1] Doob, J.L.(1953), Stochastic Processes, Wiley, New York.

[2] 王梓坤(1965),《随机过程论》,科学出版社。

[3] 戴永隆(1984),《随机点过程》,中山大学出版社。

[4] Gihman I.I.and Skorohod A.V(1974), The Theory of Stochastic Processes, Springer-Verlag, Berlin.

[5] 南京大学数学天文学系编(1961),《泛函分析》,人民教育出版社。

[6] Натансон, И.П.(1950),《实变函数论》,人民教育出版社(中译本,徐瑞云译)。

[7] 严士健,王隽骧,刘秀芳(1982),《概率论基础》,科学出版社。

[8] Фихтенгольц, Р.М.(1951),《微积分学教程》,高等教育出版社(中译本,杨弢亮等译)。

[9] 王梓坤(1976),《概率论基础及其应用》,科学出版社。

[10] 清华大学,北京大学,《计算方法》编写组(1974),《计算方法》,上册,科学出版社。

[11] Yaglom, A.M.(1962), An Introduction to the Theory of Stationary Random Fuctions, Prentice-Hall, Englewood Cliffs.

[12] Brillinger, D.R.(1981, Expanded Edition), Time Series: Data Analysis and Theory, Holden-Day, Inc.

[13] Колмогоров, А.Н.(1941),《希尔伯特空间中的平稳序列》(中译本郑绍濂等译,1962年上海科技出版社)。

[14] Akhiezer N.I. and Glazman I.M.(1961), Theory of Linear Operators in Hilbert Space, Frederick Ungar Publishing Co.

[15] С.Г.米赫林(1958),《积分方程及其应用》,商务印书馆(中译本陈传璋、卢鹤绂译)。

[16] Loève, M.(1963),Probability Theory,Van Nostrand, Princeton.

[17] Davenport, W.B.and Root, W.L.(1958), Random Signals and Noise, McGraw-Hill, New York.

[18] H. Cramér(1961), Nonstationary Stochastic Processes,Fourth Berkeley Symposium.

[19] 余家荣(1979),《复变函数》,人民教育出版社。

[20] 现代工程数学手册编委会(1985),《现代工程数学手册》,第1卷,第十八篇,华中工学院出版社。

[21] Курош А.Г.(1953),《高等代数教程》,高等教育出版社。

[22] Bucy R.S. and Moura J.M.F. edited (1983), Nonlinear Stochastic Problems, D. Reidel Publishing Co.

[23] Astrom K.J. (1970), Introduction to Stochastic Control Theory,Academic Press,New York and London.

[24] 谢衷洁,程乾生(1981),关于极大信噪比滤波,《应用数学学报》,Vol.4, No.4。

[25] Box, G.E.P. and Jenkins, G.M.(1970),Time series Analysis, Forecasting and Control, Holden-Day, San Francisco.

[26] 安鸿志,陈兆国,杜金观,潘一民(1983),《时间序列的分析与应用》,科学出版社。

[27] 中国科学院数学研究所概率组(1975),《离散时间系统滤波的数学方法》,国防工业出版社。

[28] 李卓立(1981),《实用经济计量模型与经济预测》，清华大学出版社。

[29] Rozanov,Yu.A.(1967),Stationary Random Processes, Holden-Day, San Francisco.

[30] 钱学森(1957),《工程控制论》，科学出版社。

[31] 古特金(1965),《最佳接受法理论》，科学出版社。

[32] Dwork, B.M.(1950), Detection of a pulse superimposed on fluctuation noise, PIRE, Vol.38, No.7, July.

[33] 陈兆国 (1988),《时间序列及其谱分析》，科学出版社.

[34] 汪嘉冈 (1988),《现代概率论基础》,复旦大学出版社.

[35] 韦博成 (1989),《近代非线性回归分析》，东南大学出社版。

[36] 谢衷洁(1987)，非奇异平稳序列马氏扩张的参数表示，《工程数学学报》Vol.4, 64—72.

[37] Dzhaparidze K.O. and Yaglom A.M.(1983), Spectrum parameter estimation in time series analysis, Developments in Statistics, Vol.4, AP.

[38] Shannon C.E.(1948),通信的数学理论,《信息论理论基础》，上海科技编译馆，1965.

[39] Kullback, S.(1959), Information theory and Statistics, John Wiley & Sons, Inc., New York.

[40] Haykin S. ed.(1979), Nonlinear methods of Spectral analysis, Spring-Verlag, Berlin Heidelberg.

[41] Burg J.P.(1967), Maximum Entropy spectral analysis, 37th Ann. Intern. Meeting,Soc. Exploration Geophysics, Oklahoma City.

[42] Akaike, H.(1974), A new look at the statistical model identification, I.E.E.E. Trans. Automatic

Control AC-19, 716—722.

[43] An Hong-zhi(安鸿志), Chen Zhao-guo (陈兆国) and Hannan(1982), Autocorrelation, autoregression and autoregressive approximation, The Ann. of Stat. Vol.10, 926—936.

[44] Hannan E.J. and Rissanen (1982), The recursive estimation of mixed autoregressive-moving average, Biometrika, Vol.69, 81—94.

[45] 陈兆国(1983),《ARMA 模型的定阶方法和理论》,科学院应用数学研究所概率统计咨询服务部印刷。

[46] Shibata Retei (1976), Selection of the order of an autoregressive model by Akaike's information criterion, Biometrika, Vol.63, No.1, 117—126.

[47] Spitzer F.(1956), A Combinatorial lemma and its application to probability theory. Trans. Am. Math. Soc., Vol.82, 323—339.

[48] Hannan E.J. and Quinn B.G. (1979), The determination of the order of an autoregression J.R. Statist. Soc., B. Vol.41, 190—195.

[49] Berk, K.N.(1974), Consistent autoregressive spetral estimates, Ann. Statist, Vol.2, 489—502.

[50] Priestley M.B.(1981), Spectral Analysis and Time Series Academic Press, Inc, London.

[51] Cleveland W.S. (1972), The inverse autocorrelations of a time series and their applications, Technometrics, Vol.14, 277—293.

[52] Xie Zhong-jie (谢衷洁)(1987), The Algorithm of Maximum Mutual Information for Observed Data with Additive Noise, in: Proceedings of the First World

Congress of Bernoulli Society, Vol.2, 397—399, VNU Science Press, The Netherlands.

[53] Bloomfield P. (1976), Fourier analysis of time series: An introduction, John Wiley & Sons, New York.

[54] Hannan E.J.(1980), The estimation of the order of an ARMA process, Ann. of stat., Vol.8, 1071—1081.

[55] Wan Shou-ren(王寿仁)and Chen Zhao-guo (陈兆国) (1985), Estimation of the order of ARMA model by linear procedure, Chin. Ann. of Math. Vol.6B, 53—70.

[56] Bloomfield P.(1973), An exponential model for the spectrum of a scalar time series, Biometrika, Vol. 60, 217—226.

[57] Davis, H.T., and Jones, R.H. (1968), Estimation of the innovation variances of a stationary time series, J. Amer. Statist. Assoc. Vol.63, 141—149.

[58] An Hong-zhi (安鸿志) (1982), Estimation of Prediction Error Variance, Stochastic Processes and their Applications. Vol.13, 39—43.

[59] 库洛什(1952),《高等代数教程》,高等教育出版社(柯召译).

[60] 清华大学、北京大学《计算方法》编写组(1974年),计算方法(上册),科学出版社.

[61] Schuster.A. (1898), On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena, Terr. Magn. 3, 13—41.

[62] Schuster, A.(1906), On the periodicities of sunspots, Philos. Trans. Roy. Soc., A. Vol.206, 69—100.

[63] Wang Jiagang(1983), Some Estimation Problems for Linear Model of Time Series, Chin. Ann. of Math. A, Vol.4, 19—35.

[64] 王小保(1985)，周期图的重对数律，《应用概率统计》，Vol.1，No.2，148—153.

[65] An Hong-zhi (安鸿志), Chen Zhao-guo (陈兆国) and Hannan E.J.(1983), The maximum of the periodogram, Journal of Multivariate Analysis, Vol.13, No.3.

[66] 河田龙夫(1982),《Fourier 分析》(周民强译),高等教育出版社。

[67] Anderson,T.W.(1971), Statistical analysis of time Series,New York, Wiley.

[68] U.Grenander and M. Rosenblatt(1957),《平稳时间序列的统计分析》(郑绍濂等译),上海科学技术出版社.

[69] Журбенко и.г.(1982), Спектральный анализ временных рябов, Московского университета.

[70] Hannan E. J.(1960). Time series analysis, London; Mathuen.

[71] Parzen E. (1957), On consistent estimates of the spectrum of a stationary time series, Ann. of Math. Stat., Vol.28, 329—348.

[72] Cheng Qian-sheng(程乾生)and Xie Zhong-jie(谢衷洁) (1982), Optimum high resolution window functions for spectral estimates, Progress in Cybernetics and Systems research, R. Trappl ed., Vol.8, 351—355 Hemisphere, Washington.

[73] Fisher,R.A.(1929),Tests of significance in harmonic analysis,Proc.Roy. Soc.Sec.A,Vol. 125,54—59.

[74] Whittle, P.(1952), The simultaneous estimation of *a time series harmonic components and covariance* structure, Trobajos, Estadist., Vol.3.43—57.

[75] Shimshom, M.(1971), On Fisher's test of *signi*ficance in harmonic analysis, Geophys.J. Roy. Astron.Soc., Vol.23, 373—377.

[76] He Shu-yuan (何书元) (1984), On Estimating the Hidden Periodicities in Linear Time Series Models, in: China-Japan Symposium on Statistics (Extended Abstracts of Contributed Papers), 94—97, Peking University Press, Beijing.

[77] 藤井光昭(1981),《时系列解析》,コロナ社。

[78] 谢衷洁(1985),《概率论》,人民邮电出版社。

[79] 江泽培(1980),《关于 $L^2(d\boldsymbol{F})$的可分离性》,二阶矩进修班上的报告,北京大学。

[80] 高峰(1984),《多维平稳过程的马氏扩张》,硕士论文,北京大学数学系。

[81] Yaglom A.M.(1963), Stationary Gaussian Processes Satisfying the Strong Mixing Condition and Best Predictable Functionals. Berkeley Seminar.

[82] Hannan E.J.(1970), Multiple Time Series, Wiley.

[83] Xie Zhongjie (1987), On the order discrimination of the Bloomfield model for spectral analysis of Gaussian stationary series. in: Sino-Amarican Statistical Meeting,Contributed papers, 523—526 Beijing, China.

[84] R.L.Ellis et al(1988), *In Maximum Entropy and Bayesian Methods in Applied Statistics*, Ed J.H. Justice, Cambridge University Press.